Acta Genetica et Statistica Medica

Special Volume

In association with

Edited by:

Jan A. Böök
Assistant Director of the State Institute for Human Genetics, Uppsala

Vol. IV 1953 No. 2/3

IN HONOUR OF

GUNNAR DAHLBERG

ON HIS SIXTIETH BIRTHDAY

AUGUST 22, 1953

BASEL (Switzerland) S. KARGER NEW YORK

net. Printed in Switzerland

"*Acta Genetica et Statistica Medica*" is issued quarterly. Each issue has approximately 96 pages. The annual subscription rate is Swiss frs. 48.—.

No payment is made for contributions, but 50 reprints of the article will be sent to the author free of charge. Extracopies, if desired, will be supplied at a special rate. The cost of the engravings will be borne by the publishers, provided the figures and graphs are suitable for reproduction and do not exceed a reasonable number. Otherwise the author, after due notification, will be charged with the additional cost. Articles will be printed in either English, French or German, with summaries of about 10 lines. As a rule, only original papers can be accepted.

All manuscripts should be addressed to Professor *Gunnar Dahlberg*, State Institute of Human Genetics and Race Biology, Uppsala (Sweden). Corrected proofs, review copies as well as enquiries concerning subscriptions and notices, should be sent to the publishers, *S. Karger Ltd.*, Arnold Böcklinstrasse 25, Basle (Switzerland).

«*Acta Genetica et Statistica Medica*» paraît en fascicules trimestriels d'environ 96 pages. Le prix de l'abonnement annuel est de frs. suisses 48.—.

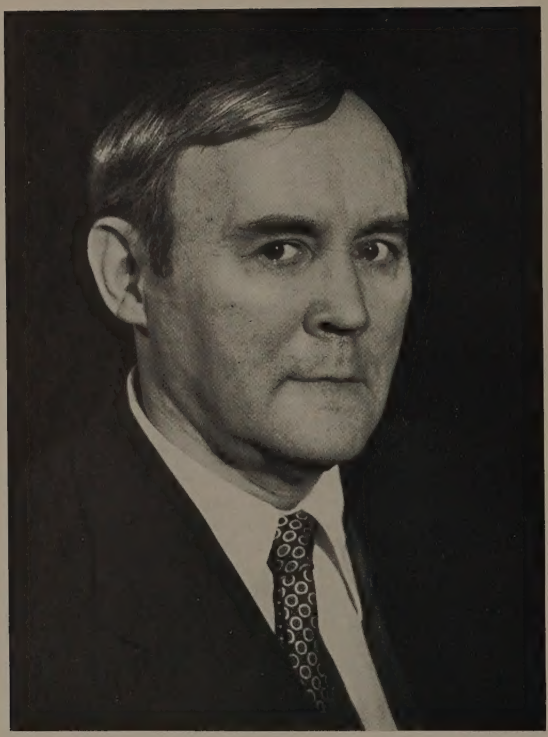
Les collaborateurs reçoivent à titre d'honoraires pour leurs travaux originaux 50 tirages à part gratuits. Les tirages à part supplémentaires seront facturés à un prix modéré. La maison d'édition se charge des frais de clichés à condition qu'elle reçoive des originaux se prêtant à la reproduction et dont le nombre ne dépasse pas la mesure strictement nécessaire. Autrement les frais supplémentaires seront, après avertissement, à la charge de l'auteur. Les travaux pourront être rédigés en langue anglaise, française ou allemande et doivent être suivis d'un court résumé d'environ 10 lignes. Ne seront acceptés en principe que des travaux originaux inédits.

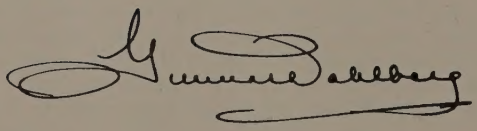
Tous les manuscrits sont à adresser au Prof. Dr. *Gunnar Dahlberg*, State Institute of Human Genetics and Race Biology, Uppsala (Suède). Les épreuves corrigées, les ouvrages à analyser, de même que toute correspondance concernant les abonnements et la publicité sont à adresser à *S. Karger S.A.* Editeurs, Arnold Böcklinstrasse 25, Bâle (Suisse).

«*Acta Genetica et Statistica Medica*» erscheint vierteljährlich in Heften von etwa 96 Seiten zum Jahresabonnementspreis von sFr. 48.—.

Mitarbeiter erhalten für ihre Originalarbeiten an Stelle eines Honorars 50 Sonderdrucke kostenfrei; weitere Separata gegen mäßige Berechnung. Die Kosten der Clichés übernimmt der Verlag, soweit reproduktionsfähige Vorlagen geliefert werden, und die Zahl der Abbildungen das notwendige Maß nicht überschreitet. Andernfalls gehen die Mehrkosten zu Lasten des Autors und werden vorher mitgeteilt. Die Arbeiten können in englischer, französischer oder deutscher Sprache eingereicht werden und sind mit einer kurzen, etwa zehnzeiligen Zusammenfassung zu versehen. Es werden grundsätzlich nur unveröffentlichte Originalarbeiten angenommen.

Alle Manuskripte sind zu richten an Prof. Dr. *Gunnar Dahlberg*, State Institute of Human Genetics and Race Biology, Uppsala (Schweden). Korrigierte Fahnen, Rezensionsexemplare sowie Zuschriften betreffend Abonnemente und Inserate sind an den Verlag *S. Karger AG.*, Arnold Böcklinstraße 25, Basel (Schweiz) zu senden.





IN HONOUR OF

GUNNAR DAHLBERG

ON HIS SIXTIETH BIRTHDAY

AUGUST 22, 1953

PUBLICATIONS BY GUNNAR DAHLBERG

Die quantitativen Beziehungen zwischen der grauen und der weißen Substanz im menschlichen Großhirn. Anat. Anz. *57*, 49–61, 1923.

The relative volumes of grey and white substance in the brain of man. Uppsala LäkFören. Förh. N. F. *28*, 351–368, 1923.

Twins and heredity. Hereditas, Lund *4*, 27–32, 1923.

A new profilometer. Uppsala LäkFören. Förh. N.F. *29*, 335–338, 1924.

Twin births and twins from a hereditary point of view. Tiden, Stockholm 1926. 296 pp.

With *J. V. Hultkrantz:* Die Verbreitung eines monohybriden Erbmerkmals in einer Population und in der Verwandtschaft von Merkmalsträgern. Arch.Rass.-u. GesBiol. *19*, 129–165, 1927.

Medizin und Mathematik. Münch. med. Wschr. *1929*, 1722.

Mittlerer Fehler und Variabilität bei Blutkörperrechnung. Acta med. scand. *70*, 103–118, 1929.

With *Gert von Sydow:* Komplikations-statistische Methoden, an Scharlach- und Diphtheriematerial dargestellt. Acta med. scand. *72*, 124–142, 1929.

Inzucht beim Menschen. Uppsala LäkFören.Förh. N.F. *34*, 849-911, 1929.

Theoretische Berechnungen über Inzucht beim Menschen. Arch. Rass.- u. GesBiol. *22*, 129-169, 1929.

Inbreeding in man. Genetics *14*, 421–454, 1929.

With *Sven Åkesson:* A theory of the uniovulation mechanism and an experimental investigation on the follicular fluid. Acta obstet. gynec. scand. *10*, 63–102, 1930.

Eine Theorie über den Uniovulationsmechanismus mit spezieller Berücksichtigung der hormonalen Wirkung des Follikulins. Klin. Wschr. *9*, 1298–1301, 1930.

Eine neue Methode zur familienstatistischen Analyse bei der Vererbungsforschung. Hereditas, Lund *14*, 73–82, 1930.

Inzucht bei Polyhybridität beim Menschen. Hereditas, Lund *14*, 83–96, 1930.

Über die Vererbung der Neigung zu Zwillingsschwangerschaft und über potentielle Fruchtbarkeit im Lichte einer Theorie der Polyovulation beim Menschen. Z. Geburtsh. Gynäk. *99*, 136–176, 1930.

Eine statistische Untersuchung über die Vererbung der Taubstummheit. Z. Konst Lehre *15*, 492–517, 1930.

Korrelationserscheinungen bei nicht erwachsenen Individuen und eine Theorie über den Wachstumsmechanismus im Hinblick auf intermittierende Umweltfaktoren. Z. Morph. Anthr. *29*, 288–306, 1931.

With *Sven Stenberg:* Eine statistische Untersuchung über die Wahrscheinlichkeit der Erkrankung an verschiedenen Psychosen und über die demographische Häufigkeit von Geisteskranken. Z. ges. Neurol. Psychiat. *133*, 447–482, 1931.

Über Bevölkerungsvermehrung. Allg. statist. Arch. *21*, 276–284, 1931.

Die Fruchtbarkeit der Geisteskranken. Z. ges. Neurol. Psychiat. *144*, 427–454, 1933.

Follicular hormone and ovulation inhibition. J. Obstet. Gynaec., Brit. Emp. *42*, 953, 1935.

On the frequency of mental deficiency. Uppsala LäkFören. Förh. N.F. *42*, 439–461, 1937.

On rare defects in human populations with particular regard to inbreeding and isolate effects. Proc. roy. Soc. Edinb. *58*, Part 2, 213–232, 1938.

Die Vererbung der allergischen Disposition. In: Fortschr. Allergie *1*, 330–351. S. Karger, Basel/New York 1939.

Alcoholism from a social point of view. A survey of conditions in Stockholm. Svenska LäkSällsk. Handl. *65*, 1–80, 1939.

On the heredity of malignant tumours. A preliminary communication. Uppsala LäkFören. Förh. N.F. *46*, 21–31, 1940.

The essentials of caries statistics and a suggestion for a standard. J. dent. Res. *19*, 479–485, 1940.

Venereal diseases in Sweden 1913 to 1937. Amer. J. Hyg. *33*, Sec. A, 51–63, 1941.

Statistical methods for medical and biological students. Allen and Unwin, London 1940. 232 pp.

With *Sten Wahlund:* The race biology of the Swedish Lapps. Part II. Anthropometrical Survey. Almqvist & Wiksell, Uppsala 1941.

A method of deciding dominance or recessivity of polymeric inheritance. Proc. 7th Int. Congr. Genet. Edinburgh 1939, p. 94, Cambridge 1941.

Rare psychological defects from the point of view of the population. Proc. 7th Int. Congr. Genet. Edinburgh 1939, p. 94–96. Cambridge 1941.

Methodik zur Unterscheidung von Erblichkeits- und Milieuvariationen mit Hilfe von Zwillingen. Hereditas, Lund, *28*, 409–428, 1942.

With *Bertil Dahlberg:* Über Karies und andere Zahnveränderungen bei Zwillingen. Uppsala LäkFören. Förh. N.F. *47*, 395–416, 1942.

With *B. Broman* and *A. Lichtenstein:* Height and weight during growth. Acta paediat., Stockh. *30*, 1–66, 1942.

The occurrence of alcoholics and their treatment. A social-statistical analysis for the town of Malmö 1929–1938. Acta med. scand. *111*, 325–358, 1942.

An analysis of the conception of race and a new method of distinguishing races. Hum. Biol. *14*, 372–385, 1942.

A theory of the nature of feelings and a suggestion for a therapy in manic-depressive mental disease. Uppsala LäkFören. Förh. N.F. *48*, 165–171, 1943.

Über genotypische Asymmetrien mit kefal-kaudaler Umkehrung bei Vererbung. Uppsala LäkFören. Förh. N.F. *48*, 195–220, 1943.

Race, reason and rubbish. A primer of race biology for the plain man. Allen and Unwin, London 1942. 240 p.

Mathematische Erblichkeitsanalyse von Populationen. Acta med. scand., supp. 148, 1943, 219 pp.

Genotypic asymmetries. Proc. roy. Soc. Edinb. Sec. B, *62*, Part 1, 20–31, 1943.

Eine Theorie über die Natur der Gefühle und ein Vorschlag zur Therapie des manisch-depressiven Irreseins. Acta psychiat., Kbh. *18*, 235-244, 1943.

Anthropometry of "tattare", a special group of vagabonds in Sweden. Uppsala LäkFören. Förh. N.F. *50*, 69–79, 1944.

With *Arthur Engel* and *Håkan Rydin:* The value of ascorbic acid as a prophylactic against "common colds". Acta med. scand. *119*, 540–561, 1944.

With *Marcus Diamant:* Inheritance of pneumatisation of the mastoid bone. Hereditas, Lund *31*, 441–456, 1945.

With *Marcus Diamant:* Hereditary character of the cellular system in the mastoid process. Acta otolaryng., Stockh. *33*, 378–389, 1945.

Nature, nurture and probability. Nature, Lond. *156*, 539, 1945.

With *Claes Grill:* Mortality in diabetes especially with regard to the treatment with insulin and diet. Acta med. scand., supp. 170, 1946, pp. 191–213.

With *E. Jorpes, S. Kallner* and *A. Lichtenstein:* Diabetes mellitus in Sweden. Acta med. scand., supp. 188, 1947. 67 pp.

Selection in human populations. Zool. Bidr., Uppsala *25*, 21–32, 1947.

Mathematical methods for population genetics. S. Karger, Basle/New York 1947. 182 pp.

Vererbung und Rasse. Phönix-Verlag Christen & Co., Hamburg 1947.

Genetics of human populations. In: Advances in genetics *2*, 67–98, Academic Press Inc., New York 1948.

The Swedish bed-state. Brit. J. soc. Med. *2*, 53–54, 1948.

With *A. B. Maunsbach:* The eruption of the permanent teeth in the normal population of Sweden. Acta genet. *1*, 77–91, 1948.

Environment, inheritance and random variation with special reference to investigations on twins. Acta genet. *1*, 104–114, 1948.

A new method in crime statistics, applied to the population of Sweden. J. crim. Law, Northwestern University *39*, 327–341, 1948.

With *Elvir Lander:* Size and form of the foot in men. Acta genet. *1*, 115–162, 1948/49.

Do parents want boys or girls? Acta genet. *1*, 163–167, 1948/49.

Notes on the conception of type. Acta genet. *1*, 174–178, 1948/49.

The frequency of orthostatic anaemia at different ages. Acta genet. *1*, 179–190, 1948/49.

Suicide, alcohol and war. Acta genet. *1*, 191–198, 1948/49.

With *F. Sundelin:* Rheumatoid arthritis and the function of the joints. Acta med. scand. *135*, 40–46, 1949.

Fettleibigkeit and Diabetes. Acta paediat., Stockh. *38*, 91–107, 1949.

The thymus and obesity. Acta genet. *1*, 199–218, 1949.

Blood sinking reaction from a statistical view-point in normal men and women. Acta genet. *1*, 267–276, 1949.

The frequency of cases of disease. Acta genet. *1*, 277–285, 1949.

What does normal mean? Acta genet. *1*, 286–294, 1949.

Mortality from tuberculosis in some countries. Brit. J. soc. Med. *4*, 220–227, 1949.

Les jumeaux et l'ovulation. Rev. sci., Paris *87*, 141–147, 1949.

Standard error and medicine. Acta genet. *1*, 313–321, 1950.

The prognosis of disease. Acta genet. *1*, 322–328, 1950.

Venereal disease and prostitution. Acta genet. *1*, 329–342, 1950.

Obesity and diabetes. Acta genet. *1*, 343–354, 1950.

With *Greta Haglund* and *Nils G. Richtnér:* The therapeutic significance of tonsillectomy in polyarthritis. Acta otolaryng., Stockh. *38*, 306–311, 1950.

Thymus und Fettleibigkeit bei Mäusen. In: Moderne Biologie, Festschrift zum 60. Geburtstag von Hans Nachtsheim. Peters, Berlin 1950, S. 23–28.

With *W. Nordlöw:* Genetics of convergent strabismus. Acta genet. *2*, 1–14, 1951.

Mental deficiency. Acta genet. *2*, 15–29, 1951.

A note on drinking and drunkards in Sweden. Acta genet. *2*, 36–41, 1951.

With *Ivan Grubb:* Chronic rheumatic arthritis and housing conditions. Acta genet. *2*, 42–56, 1951.

An explanation of twins. Sci. Amer. *184*, 48–51, 1951.

A note on the relative death rate. Acta genet. *2*, 173–174, 1951.

The primary sex ratio and its ratio at birth. Acta genet. *2*, 245–251, 1951.

Eine Methode zur Messung von Rassenunterschieden. Metron, Roma *16*, N. 3–4, 1951.

Marriage and divorce risks. Acta genet. *2*, 318–330, 1951.

Sociology, race and genetics. Proc. 14th Int. Congr. Sociology Rome 1950, Vol. II, pp. 803–811, Roma 1952.

Die Tendenz zu Zwillingsgeburten. Acta Ge. Me. Ge. *1*, 80–88, 1952.

The mortality rate of cancer yesterday and today. Acta genet. *3*, 61–69, 1952.

A note on the dilemma of human genetics. Acta genet. *3*, 69–72, 1952.

The Swedish jews. Acta genet. *3*, 106–117, 1952.

Genetic investigations in different populations. Acta genet. *3*, 117–142, 1952.

Death rates in Swedish provinces and the effect of isolates. Acta genet. *3*, 143–158, 1952.

With *B. Josephson:* Variations in the cell-content and chemical composition of the human blood due to age, sex and season. Scand. J. clin. Lab. Invest. *4*, 216–236, 1952.

With *B. Josephson* and *G. Tötterman:* On the effect of nutrition upon the blood count and chemical composition of the blood. Scand. J. clin. Lab. Invest. *4*, 237–241, 1952.

Notes on differentiated population increments. Acta genet. *4*, 101–117, 1953.

Biometric evaluation of findings. In A. Sorsby: Clinical Genetics, 83–100. Butterworths, London 1953.

Dr. Dahlberg has also published a large number of papers in Swedish periodicals, articles in newspapers and a number of books in Swedish. Among the books the following should be mentioned:

A medical encyclopedia, published for the Swedish Red Cross («Svenska röda korsets medicinska uppslagsbok»). 12th ed. AB Svensk Litteratur, Stockholm 1952, 1230 pp.

A general encyclopedia («Tidens Lexikon»). 5th ed. Tiden, Stockholm 1950. 3270 pp.

Diseases and society. («Sjukdomarna och samhället».) I-II. Bonniers, Stockholm 1934–37. 475 pp.

With *Herbert Tingsten:* A political encyclopedia. («Svensk Politisk Uppslagsbok».) AB Svensk Litteratur, Stockholm 1937. 397 pp.

«Vi och vår värld.» A scientific review. AB Svensk Litteratur, Stockholm 1941. 868 pp.

«Dit min tanke nått.» Biological and sociological essays. Bonniers, Stockholm 1943. 326 pp.

The future society. («Det framtida samhället».) Tiden, Stockholm 1944. 163 pp.

Mother and child. («Mor och Barn. En bok om barnens vård och fostran, utgiven av Svenska röda korset.») 4th ed. AB Svensk Litteratur, Stockholm 1946. 479 pp.

Normal and pathological genetics. («Normalt och sjukt arv».) Bonniers, Stockholm 1946. 350 pp.

«Infall och utfall.» Social and biological essays. Tiden, Stockholm 1946. 251 pp.

With *Ivar Johansson:* A Swedish Agricultural Encyclopedia. («Svenskt Lantbrukslexikon».) I-II. 2nd ed. AB Svensk Litteratur, Stockholm 1947. 1136 pp.

A popular medical textbook. («Tidens läkarbok».) 3rd ed. Tiden, Stockholm 1948. 526 pp.

With *S. A. Gammeltoft:* A book on the woman. («Boken om kvinnan».) I-II. Örnkrona, Stockholm 1951. 1028 pp.

CONTENTS

With the Assistance of a Grant from Hässle Pharmaceutical Company, Hässleholm, Sweden

Böök, J. A.: Acta genet. 4, 133–139, 1953

From the State Institute for Human Genetics, Uppsala, Sweden
(Head: Professor Gunnar Dahlberg, M. D. LL. D.)

SCHIZOPHRENIA AS A GENE MUTATION

By J. A. BÖÖK

In a recent paper (*Böök* [1953]), I have scrutinized the accumulated facts on which the genetic theory of schizophrenia was founded. It was concluded that this theory is scientifically sound and that some kind of gene action would be the *basic* etiologic agent for the majority of schizophrenic psychoses. During the past few years, the genetic theory of schizophrenia has become more and more accepted by psychiatrists, with the notable exception of those who base their etiological concepts on psychoanalytic hypotheses.

In the following discussion it will be assumed that the genetic theory of schizophrenia is by and large correct. This is permissible as the criticism so far offered by opponents has not been of such a kind as to invalidate or even seriously question the theory. It is another matter that psychiatric geneticists still disagree about the more *detailed nature* of the postulated genetic factor, its type of transmission and so forth.

I have studied the occurrence of schizophrenia in a North Swedish isolate of about 9,000 individuals during the period of 1902–1949. Only those *results* which have a direct connection with the discussion of gene mutations can be mentioned in this paper. For detailed information about the data the reader is referred to *Böök* [1953].

The schizophrenic psychoses observed in this population displayed a rather uniform clinical picture, interpreted as predominantly catatonic with a marked tendency to periodicity. The calculation of the general morbid risk gave a value of 3 per cent which is the highest figure so far found in any population. The closer genetic analysis favoured the hypothesis of a simple dominant gene with complete penetrance in homozygotes and about 20 per cent penetrance in heterozygotes. The gene frequency (d) was estimated at 7 per cent.

A special analysis of the reproductive capacity of schizophrenic subjects resulted in an estimated fitness (f) of 70 per cent as compared with non-neuropsychiatric subjects.

It is well known that the reproductive capacity of schizophrenic subjects is below normal (*Essen-Möller* [1935]). Furthermore, we have no indications that the incidence of schizophrenia has decreased or increased appreciably during the admittedly rather few generations about which we have more or less scanty information. It is more likely that it is maintained at a state of relative equilibrium, the level of which might, of course, be different in different populations. The logical sequence of this is that a certain number of mutations would be needed to counteract the losses due to selection.

Taking the North Swedish population which was studied in detail and assuming the validity of the *Hardy-Weinberg* equation within the isolate, we have per generation $d^2 + 2drp$ (p = penetrance) individuals who will get the disease, assuming that they survive long enough, and $2dr(1-p)+r^2$ who will escape. In a population of N the number of eliminated genes per generation is

$$\frac{2d^2N(1-f) + 2drpN(1-f)}{2N}$$

where f, as mentioned above is the reproductive fitness. This equals the mutation rate if equilibrium is to be maintained, i.e. if the allele frequencies d and r remain unchanged. That such an equilibrium prevails in this population is, of course, questionable. In any case, the figure should be a minimum. Simplifying the equation, the mutation rate will be

$$u = (1-f)(d^2 + drp) \qquad (1)$$

Substituting the observed and derived values of f, d, r and p, we obtain a mutation rate of 5×10^{-3} genes per generation. If this figure is correct only in so far as its order of magnitude is concerned, it represents a very high mutability. It is well known that all estimates of mutation rates of human genes are extremely uncertain and so far we must be content with what comes next to guesswork, although with a certain amount of sense behind it. The difficulties involved in calculations of this kind have been submitted to rather detailed discussions by *Haldane* [1949], *Muller* [1950] and *Neel* [1952].

Previous estimates of the mutability of other human genes have mostly resulted in figures about 10^{-5} (*Haldane* [1949], *Böök* [1952]). It is, however, noteworthy that in a few cases much higher estimates

have been made. So, for instance, *Reed*'s figures of 5.1×10^{-4} for the Rh loci and $0.7-1 \times 10^{-3}$ for cystic fibrosis of the pancreas (cf. *Neel* [1952]) as well as the figure of 1×10^{-3} for the sickle cell disease (*Neel* [1951]) should be mentioned.

Previous estimates of mutation rates in man so far concerned, almost exclusively, very rare genetic disorders. It is possible that conditions like epiloia, chondrodystrophy, retinoblastoma, albinism, infantile amaurotic idiocy and hemophilia are due to genes of relatively low mutability, although the estimated mutation rates (all about 10^{-5}) average higher values than most of the known rates in *Drosophila*. We know, e.g. from the experimental work of *Stadler* [1942], that some genes (of the same organism, in this case maize) are distinctly more mutable than others. It has also been pointed out by *Dobzhansky* [1951] that the genes in higher animals and plants apparently mutate more frequently, per generation, than those in microorganisms. This could be due to the appreciable differences of the duration of generations. Thus one should expect that the mutation rate per chromosome per unit time would be approximately the same. However, as *Muller* [1950] emphasized, the human germ cell lineage includes two or more times as many cell divisions as that of *Drosophila* (where it is known that the mutations are concentrated to these nuclear phases). This together with the higher temperature in the human should more likely tend to give higher mutation rates as compared with the fly.

Several reasons could be given to explain the calculated mutation rates in man as being either too low or too high. They would be too high if the traits so far studied were not distinct clinical *and* genetical entities but a mixture of genetic entities with similar or identical phenotypical expression. They would also be too high if unmanifested gene carriers displayed a positive selective value. If these latter had a fitness *below* unity, the calculated rates would be too low. Especially in regard to *complete recessivity* the fitness of the heterozygotes is important because they are generally so much more numerous than the homozygotes. At present we know almost nothing about these things. However, we know a little more about some important trends of the genetics of human populations during the last few centuries. *Dahlberg* [1938] analysed the effect of enlarging the size of isolates, a process that has been going on parallel with technical developments. This means that old equilibria in smaller isolates are disturbed and slowly become replaced by new ones in larger isolates. It also means

a decrease of rare recessive homozygotes, since such individuals have occurred in some isolates but not in others. Consequently as the breaking up of small isolates has been a relatively recent phenomenon and the establishment of new equilibria might be a matter of some hundred years, the incidence of homozygotes which we observe now in the enlarged isolates does not correspond to what should be expected on the basis of the *Hardy-Weinberg* law but is much too low. Thus, the calculated mutation rates as based on the number of homozygotes must also be too low. It seems probable that those mutation rates that can be estimated in small or relatively small geographical isolates come closer to the true figures. It is more likely that selection and mutation pressure are closer to the equilibrium point here.

As an example I could mention a special type of mental defect, genetic spastic oligophrenia (*Böök* [1953]), which occurred in the above-mentioned North Swedish population with a prevalence rate of 0.002. As the reproductive fitness in this case equals 0, the estimated mutation rate could be as high as 2×10^{-3}, i.e. of the same order of magnitude as was estimated for schizophrenia in the same region.

If mutation rates of this magnitude occur in man and if they are not rare exceptions, it would imply a much greater evolutionary plasticity than hitherto imagined.

On a *priori* grounds it is not at all unlikely that man posesses a number of unstable i.e. highly mutable genes which in case the mutated genes display a low penetrance or else cause only moderately reduced adaptability might be difficult to disclose. In so far as mental traits are concerned, this possibility will remain an object of prime importance for future research, although the discussion today must be of a speculative nature. The last decade of research in medical genetics has shown, I believe beyond doubt, that mutations as a cause of human diseases can no longer be considered negligible.

Returning to the mutation rate of schizophrenia as estimated at 5×10^{-3} in this population, a few consequences seem worthy of comment. The implication is that about one out of 100 newborns would carry a newly mutated gene. As such an individual would be heterozygous, he would be expected to manifest a schizophrenic psychosis with the probability of about 20 per cent (equalling the penetrance). Thus of 1,000 newborns, 2 would be predetermined schizophrenics or, expressed otherwise, the general morbid risk of schizophrenia due to new mutations would be about 2 per 1,000. As

we have calculated a general morbid risk of 30 per 1,000 based on the actually observed number, this also implies that 2/30 or between 6 and 7 per cent of the observed cases would be new mutations. As such, the incidence of schizophrenia among their parents and sibs would have no connection with the fact that these latter individuals are related to a schizophrenic *propositus*, but correspond to the general risk. Apart from the important possibility that an appreciable fraction of the schizophrenics could be new mutations, a certain amount of error would be introduced in the calculations of the genetic analysis. This error would not be large enough to invalidate the hypothesis of transmission as put forward here. Although there is at present no way to make corrections, certain consequences could be anticipated.

In so far as the morbid risk of siblings with one schizophrenic parent is concerned, the new mutations would make no difference. However, a certain number of schizophrenic *propositi* from two non-schizophrenic parents, possibly as much as 10 per cent, would be new mutants. This would mean that the calculated morbid risk for this combination was somewhat too low since a number of siblings were included who, as relatives of the mutants, should display no increased risk. This is interesting in view of the fact that the latter risk was actually found to be somewhat lower although not statistically different from the risk for siblings from one schizophrenic parent (9 per cent against 12). A similar and not significant difference has also been found by other workers e.g. *Kallmann* [1938]. It should be mentioned, however, that *Kallmann* has entirely different ideas about the genetic mechanism (for further discussion see *Böök* [1953]) and that the type of schizophrenia with which we are concerned in the present population might be due to a different gene.

Some further consequences are interesting. Assume a change in the biologic structure of the population so that the reproductive fitness decreases to 0.25 for schizophrenic subjects. This figure is in agreement with some of the estimates made by *Essen-Möller* [1935] for a large South German population. The mutation rate is assumed constant at 5×10^{-3}, likewise the penetrance (p) for heterozygotes at 0.2. Substituting $(1-f) = 0.75$ in equation (1) and solving it in regard to d we obtain a figure of 0.03. This means that the general morbid risk now would be reduced to approximately 1 per cent. Furthermore about 20 per cent of individuals predetermined to get the disease would be new mutations. Consequently, starting with schizophrenic

propositi in such a population, the calculated morbid risks for their parents and siblings from two non-schizophrenic parents would be lower than in the preceding example. Under such circumstances schizophrenia would virtually behave more like a recessive trait.

It may justifiably be asked if the high prevalence rate in the present population could be due to the relatively high reproductive fitness of schizophrenic subjects here as compared to other populations. Such an explanation could, of course, be correct only if schizophrenia were a genetic entity, a question which cannot yet be answered.

The etiology of schizophrenia has become one of the most important problems not only of psychiatry and medicine but of our present day culture. Being aware of its extremely complicated nature, the above paragraphs have been offered simply as a discussion of some more or less likely possibilities. As I have found the genetic theory to be the best adapted integer of the coincidences of facts now available, it seemed justified to present here some of the consequences of this theory under the specific conditions which prevailed in the studied population. Even if the interpretations which have been offered here might prove to be completely mistaken, they hold possibilities of fundamental importance and thus invite to further studies.

Summary.

On the basis of an extensive study of schizophrenia in a North Swedish population, the writer discusses the possibilities that the type of psychosis prevalent in this region, interpreted as primarily due to an autosomal dominant gene, might rather often originate through new mutations. With all reservations for the insecurities involved in such calculations, the mutation rate was estimated at 5×10^{-3}, thus indicating a very mutable gene. Some of the consequences for the interpretations of the genetics of schizophrenia are also discussed very briefly.

Résumé.

En se basant sur des études approfondies concernant la schizophrénie dans les populations du nord de la Suède, l'auteur se demande si le genre de psychose, qui est fréquent dans ces régions et que l'on considère comme dépendant d'un gène dominant autosomique, ne pourrait pas assez souvent être provoqué par de nouvelles mutations. En formulant des réserves sur le caractère d'incertitude que présente ce genre de calcul, la fréquence des mutations est estimée à 5×10^{-3}, ce qui indiquerait un gène très mutable. On traite également en résumé les conséquences que cette conclusion entrainerait pour les conditions d'hérédité de la schizophrénie.

Zusammenfassung.

Unter Zugrundelegung umfangreicher Studien über Schizophrenie in nordschwedischer Population diskutiert der Verfasser die Möglichkeit, daß derjenige Typ

von Psychose, welcher in diesem Gebiete auftritt und welcher als primär bedingt durch ein autosomal dominantes Gen erklärt wurde, häufig durch neue Mutationen entstehen könne. Mit vollster Reservation für die Unsicherheitsmomente, welche derartigen Berechnungen anhaften, wird die Mutationsfrequenz auf 5×10^{-3} geschätzt, welches demgemäß ein sehr mutables Gen bedeuten würde. Eine Reihe der Konsequenzen, welche dies für die Deutung der Erblichkeitsverhältnisse der Schizophrenie mit sich führen würde, wird hierauf kurz besprochen.

REFERENCES

Böök, J. A.: J. Genet. hum. *1*, 24–26, 1952; A Genetic and Neuropsychiatric Investigation of a North Swedish Population, with special regard to schizophrenia and mental deficiency. Part I. Psychoses. Acta genet. *4*, 1–100, 1953. – *Dahlberg, G.:* Proc. Roy. Soc. Edinb. *58*, 213–232, 1938. – *Dobzhansky, Th.:* Genetics and the Origin of Species, 3rd ed. 364 p. Columbia Univ. Press., New York 1951. *Essen-Möller, E.:* Acta psychiat., Kbh. Suppl. *8*, 314, 1935. – *Haldane, J. B. S.:* The rate of mutation of human genes. Proc. 8th int. Congr. Genet., Hereditas. Suppl. 267–273, 1949. – *Kallmann, F.:* The Genetics of Schizophrenia. 291 p. J. J. Augustin Publ. New York 1938. – *Muller, H. J.:* Amer. J. Hum. Genet. *2*, 111–176, 1950. – *Neel, J. V.:* Cold Spr. Harb. Symp. Quant. Biol. *15*, 141–158, 1951; Amer. Nat. *86*, 129–144, 1952. – *Stadler, L. J.:* Spragg Mem. Lect. *3*, 3–15, 1942.

Dunn, L. C.: Acta genet. *4*, 139–151, 1953

Institute for the Study of Human Variation, Columbia University, New York City
With an Appendix by Timothy Prout, Columbia University

VARIATIONS IN THE SEGREGATION RATIO AS CAUSES OF VARIATIONS IN GENE FREQUENCY

By L. C. DUNN

Among the causes responsible for changes in the relative frequency in a population of the members A and a of a pair of alleles, the most frequently cited are 1) mutation of $A \rightarrow a$ and $a \rightarrow A$ at unequal rates; 2) selection, or inequality between the different genotypes in the production of surviving progeny; 3) differences between the genotypes in the rate at which the different genotypes enter or leave the population by migration; 4) chance fluctuations in the frequency of A and a due to sampling in small populations, leading to or toward fixation of AA or aa.

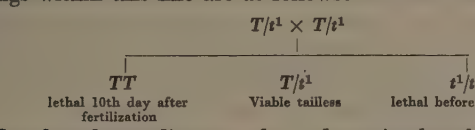
It can hardly be doubted that these factors, singly or in concert, are those most likely to be responsible for those disturbances in the genetic equilibrium of cross-breeding populations which lead to evolutionary changes. Nevertheless, a few cases are known in which variation of another kind must be taken into account. These are cases in which the disturbing factor is an inequality in the frequency of gametes *A* and *a* produced by heterozygotes *Aa*. Mendelian heredity and its corollary, *Hardy-Weinberg* equilibrium in panmictic populations, assume the equality of such *A* and *a* gametes as a matter of course, and the assumption is generally justified by direct evidence and by success in application. But the rule is not universal, as its infringement in the case of the sex-linked "sex-ratio" mutation in several species of Drosophila clearly showed (*Morgan*, *Bridges*, and *Sturtevant* [1925]; *Gershenson* [1928]). Males with this mutation produce chiefly daughters. In *D. pseudoobscura* this is due to an irregularity in spermatogenesis of such males by which the X-chromosome divides twice while the Y-chromosome is ordinarily not included on the spindle, so that nearly all the sperm receive an X (with the mutation) and very few a Y (*Sturtevant* and *Dobzhansky* [1936]). This great excess of mutant sperm should lead to a rapid spread of the sex-ratio mutation. Indeed it should soon become fixed in any population in which it occurs, and such a population would then consist almost entirely of females. Its spread, however, is opposed by the disadvantageous effects of the mutation in homozygous females and in males, so that, as *Wallace* [1948] has shown, its frequency in a confined population declines when it is in competition with its normal allele.

The discovery of a series of alleles in the house mouse, most of which are transmitted through the male in such a way as to constitute valid exceptions to the expected 1:1 segregation ratio, has directed attention to the effects which deviations of this kind may have on the evolutionary fate of such alleles in mammals. The results from studies of mouse populations suggest that deviations from normal segregation ratios should be considered in those cases in which certain alleles in human populations are found to have unexpectedly high values.

Inheritance of t-alleles in the house mouse.

In the house mouse an unusual opportunity for detecting mutative changes at or near one locus was afforded by the discovery

that certain inbred lines, breeding true to taillessness, constitute in fact balanced lethal systems (*Chesley* and *Dunn* [1936]; *Dunn* [1937]). One such line and its derivatives have been extensively studied (*Dunn* and *Gluecksohn-Waelsch* [1953]) with the following results. The line is constituted by two lethals T and t^1 which show no recombination and hence can be treated as alleles. The usual results of matings within this line are as follows:



In the above diagram the only animals which survive are heterozygotes which are tailless. The same results are observed with T/t^0 tailless, t^0 being an allele of independent origin.

As exceptions viable animals with normal tails have been found at birth. Eleven such exceptions were noted among 3500 offspring observed from matings of $T/t^1 \times T/t^1$. Seven of these were analyzed in extensive breeding experiments and in each case the normal exception was shown to be t^1/t^x, t^x being a new allele of T which produces a tailless phenotype in combination with T, but normal tail in combination with t^1. Of the seven new alleles discovered in this way, four (t^3, t^7, t^8, t^{11}) give viable normal tailed homozygotes, three (t^4, t^9, t^{12}) are lethal. None show recombination with T.

In outcrosses of these and other t-alleles to animals with normal alleles at this locus a difference in the outcome of reciprocal crosses has been noted. Thus, T/t^x♀$\times +/+$♂ produces equal numbers of $T/+$ (short-tailed Brachyuric) and $t^x/+$ (normal-tailed) offspring; while T/t^x ♂$\times +/+$♀ usually produces a departure from this ratio. In the cases of t^0, t^1, t^2, t^3, t^9, t^{12}, heterozygous males produce more progeny with the t^x allele than with the allele T or $+$, from which it is segregating; while with one allele (t^4) the excess is in the opposite direction; t^7, t^8, t^{11} show no consistent direction of departure.

In the case of some alleles the excess of one type of progeny is very great. For example, T/t^1 males $\times ++$ females (all from the same inbred Bagg albino line) gave 647 normal and 88 Brachy; while T/t^{12} ♂ $\times ++$ ♀ gave 187 normal and 13 Brachy. In some of the most extreme cases a large sample of such F_1 normal progeny has been tested and shown to be $+/t^x$. The cause of the altered ratio is thus the production of T and t^x (or $+$ and t^x) sperms in unequal members, the t^x sperm usually being in excess. The abnormality

in spermatogenesis responsible for this has been sought by both genetical and cytological methods but has not been identified.

Since the above cases of altered segregation ratios occurred in inbred laboratory stocks containing other mutant genes, no attempt to assess their probable effects on gene frequency in populations could be undertaken until the stocks to be compared were rendered isogenic at other loci. While this was being done, several males taken from wild populations were tested by crossing with Brachy ($T/+$) females and a surprising result obtained. In most of the wild populations tested from different parts of the United States (*Dunn* and *Morgan* [1953]), heterozygotes $+/t^w$ were found (t^w being a *t*-allele occurring in the wild) and in all cases $+/t^w$ males by $T/+$ females yielded many more T/t^w (tailless) than $T/+$ (Brachy) offspring, indicating the same kind of excess of *t* sperm which had been found in the laboratory stocks. The first results of testing wild males are shown in Table 1.

Table 1. Results from testing $+/t^{w1}$ males from wild populations by short-tailed (*Brachy*) females $T/+$.

Male No.		Source	Progeny: Normal $+/+$; $+/t^w$	Brachy $T/+$	Tailless T/t^w
4922	$+/t^{w1}$	New York 1	23	2	20
4927	$+/t^{w1}$	New York 1	29	5	20
4929	$+/t^{w1}$	New York 1	14	2	7
4930	$+/t^{w1}$	New York 1	6	–	4
4928	$+/t^{w2}$	New York 1	80	3	54
6380	$+/t^{w2}$	New York 1	47	4	38
6089	$+/t^{w4}$	Wisconsin 3	16	1	10
N.Y.2	$+/t^{w5}$	New York 2	8	3	12

In order to test the assumption that the ratio of different tail types among the progeny is due to the ratio between $+$ and t^w gametes of the wild males, all normal-tailed offspring of wild male 4928 were mated with $T/+$. Of the 46 which survived and bred, 42 (17 ♂, 25 ♀) gave tailless progeny and proved to be $+/t^w$; 4 (1 ♂, 3 ♀) gave only normal and Brachy and proved to be $+/+$. The ratio of $+/t^w$ to $+/+$ among these normal progeny is thus far from equality and similar to the ratio of tailless (T/t^w) to Brachy ($T/+$) in F_1 which we may thus take to express the ratio of t^w to $+$ gametes produced by the wild male. The total ratio was thus 96 t^w to 7$+$. It can be safely assumed that each wild male tested produced many more t^w

than + gametes, and thus that t^w enjoys a great numerical advantage in spermatogenesis relative to +.

The questions then arose whether this advantage was inherent in the t^w allele or was influenced by the balance of the residual genotype of wild males; whether different t^w alleles segregated in similar or different ratios; and whether the segregation of t^w from T/t^w heterozygous males was similarly abnormal. Answers to these questions were obtained from the experiment summarized in Table 2.

Table 2. Results of testing T/t^w males individually by +/+ females.

Male No.		Offspring Normal +/t^n	Offspring Brachy +/T	Total Normal	Total Brachy
6267	T/t^{w1}	49	3		
6598	T/t^{w1}	63	9		
6599	T/t^{w1}	22	8		
7446	T/t^{w1}	14	2	148	22
5823	T/t^{w2}	39	1		
6701	T/t^{w2}	33	–		
6779	T/t^{w2}	67	1		
7202	T/t^{w2}	24	–	163	2
6006	T/t^{w3}	85	1		
6623	T/t^{w3}	63	–		
6910	T/t^{w3}	51	2		
6914	T/t^{w3}	52	–		
6915	T/t^{w3}	27	–	278	3
6708	T/t^{w4}	44	–		
6710	T/t^{w4}	25	–	69	–
7427	T/t^{w5}	6	–		
7436	T/t^{w5}	14	–		
7655	T/t^{w5}	12	1		
7705	T/t^{w5}	5	3	37	4
Total				695	31
Per cent				95.7	4.3

It is evident that males carrying t^w alleles (in combination with T) produce two classes of offspring in very unequal proportions. The dominant allele, T, appears in less than five per cent of the offspring, whereas when transmitted by $T/+$ males it appears in the expected 50 per cent. A sample of the normal-tailed offspring from each T/t^w

male is being tested. To date no exceptions have been found to the rule that such offspring are $+/t^w$ (earlier tests of males similarly derived from T/t^0 males showed that all normal-tailed offspring were $+/t^0$). Hence we may take it that the progeny ratio of normal to Brachy expresses in general the ratio of t^w to $+$ gametes formed by T/t^w males. All tested males agree in transmitting t^w in the great majority of sperm. The advantage of t^w thus probably inheres chiefly in its own effect on spermatogenesis. The sperm ratio of t^{w1} is somewhat lower than that of the other t^w alleles and this difference is significant indicating minor differences between t^w alleles. This had also been shown for t alleles arising as mutations from balanced lethal stocks in the laboratory (*Dunn and Gluecksohn-Waelsch* [1953]).

The spread of t^w alleles in natural populations.

The result of chief interest from the above experiments is that t^w alleles found in the wild mice are transmitted in spermatogenesis in far higher frequency than either the T or $+$ alleles. This alteration of the normal segregation ratio should thus result in the replacement of other alleles by t^w in all populations in which t^w alleles occur, unless this is opposed by other circumstances. The chief hindrance to the spread of t^w alleles is the fact that of the five studied four $t^{(w1, 3, 4, 5)}$ are lethal and one (t^{w2}) although viable produces sterility in homozygous males. Males containing two different t-alleles (t^{w1}/t^{w2}, t^{w1}/t^{w4}) are also viable and sterile.

It is therefore necessary to determine the effects on equilibrium in natural populations of modifications in the segregation ratio as compared with the effects of selection. No theory has hitherto been developed for this since the segregation ratio $A:a$ from A/a has usually been assumed to be .5. Consequently Mr. *Timothy Prout*, at my suggestion, undertook to formulate a theory which is described by him in the appendix to this paper. I am greatly indebted to him for permission to include this formulation and to quote from it.

The results relevant to the present case are that a recessive deleterious mutation which has a segregation-ratio ($m > .5$) relative to its allele, as the t^w alleles have, can spread through a population to reach equilibrium or fixation. Even when the mutation is completely lethal in homozygotes, as in the cases of t^{w1}, t^{w3}, t^{w4}, t^{w5}, it can still come to equilibrium if m be sufficiently high. The general condition for equilibrium is given by

$$q = \frac{r + 2m - 1}{r + s}$$

where r and s are the selection coefficients for the + and t^w homozygotes respectively. In the case of alleles like t^w, which segregate normally in females ($m♀ = .5$) but abnormally in males ($m♂ = 95$), the value of m for equilibrium computations must be $(m♀ + m♂)/2$. If $r = 0$ (no selection against ++), $s = 1$ (complete selection against t^w) and $m = .75$, which is close to that found in the first populations tested (Table 2), then $\hat{q}t^w = .5$. This gametic frequency should produce zygotic frequencies:

.25 +/+ .5 +/t^w .25 t^w/t^w (dead embryos).

Thus in a population in which these conditions obtain, two-thirds of the individuals should be heterozygous for the lethal; and these proportions should be stable.

One of the populations tested (New York 1) had been bred in captivity as a small colony for six years[1] being reduced to a few animals each summer so that considerable inbreeding occured. We tested 10 males and found 2 sterile, $1+/t^{w2}$(t^{w2}/t^{w2} is male-sterile), $4+/t^{w1}$ (t^{w1}/t^{w1} is lethal), and 3 +/+. The t^w alleles had persisted in the majority in spite of lethality, and sterility and inbreeding, probably because of the segregation ratio advantage.

Under such conditions different wild populations could be in equilibrium for different t^w lethals which would be maintained by segregation ratios against the pressure of selection.

According to *Prout's* analysis, it is not until deleterious dominant effects are also ascribed to the lethals that we encounter conditions capable of counteracting the spread of such alleles. The fact that t^w alleles have not been found in all populations tested (actually none found in small populations from Vermont, Maine, and Kansas; t^w alleles found in New York 1, New York 2, Wisconsin 3, Connecticut, Florida), nor in high frequency in all populations in which one t^w allele has been detected, suggests that some t^w alleles may have deleterious (partly dominant) effects on heterozygotes. Experiments have been started to test this hypothesis.

---

[1] By Dr. *Howard A. Schneider* of the Rockefeller Institute for Medical Research of New York.

Suggested application to human populations

In several cases, alleles with deleterious recessive effects are found in human populations with frequencies above the expected equilibrium values. This is true, for example, of the sickling alleles (*Neel* [1951, 1952]; *Hiernaux* [1952]) and of thalassemia (*Silvestroni* [1950]). Two explanations have been suggested: first, that mutations to the disadvantageous allele occur at a high rate ($\hat{q} = \sqrt{u}$, where u is the mutation rate), or to a selective advantage of the heterozygote. The first assumption leads to mutation rates higher than any observed in mammals and hence unlikely; the second calls for greater heterozygote advantage than has been found in any such case. A relatively minor segregation-ratio advantage of the deleterious allele would, as *Prout's* analysis shows, suffice for maintenance of an equilibrium without further assumptions. Although no evidence of such segregation-ratio differences have been found in human populations, they have generally not been looked for. It would not be unreasonable to look for such abnormalities to occur during meiosis at such complex loci as Rh and MNS, where the multiple antigenic specificities suggest duplication of genetic material with suppression of crossing over or other local interference with normal meiotic events.

Summary.

Evidence is presented showing that recessive deleterious alleles (four lethals and one male sterile) found in wild populations of house mice enjoy a great segregation-ratio advantage. About 95 per cent of the sperm of heterozygous males ($+/t^w$) transmit the deleterious allele. It is shown that alleles with segregation-ratio advantage can be retained in a population at equilibrium frequencies even against adverse selection pressure. This proposition is generalized in the appendix, and equilibrium values for variations in m (segregation ratio), r and s (selection coefficients of the respective alleles) are given. The general expression for gene frequency at equilibrium is

$$\hat{q} = \frac{r + 2m - 1}{s + r}$$

It is suggested that deleterious alleles and others at complex loci which have high equilibrium values in human populations may be favored by segregation ratio advantage.

Résumé.

L'investigation montre que des allèles récessifs tarés (4 gènes léthaux et 1 gène déterminant la stérilité du mâle) qu'on a trouvés dans des populations

sauvages de souris ordinaires, montrent un taux de ségrégation très positif. Environ 95 % des spermazoïdes des mâles hétérozygotes ($+/t^w$) transmettent l'allèle taré. L'auteur montre que de tels allèles ayant un taux de ségrégation positif peuvent être maintenus dans une population en équilibre de fréquence malgré une pression de sélection inverse. Cette interprétation est exposée sous une forme générale dans l'appendice ajouté par *T. Prout* et les valeurs d'équilibre pour des variations différentes de m (le taux de ségrégation) ainsi que celles de r et de s (les coefficients de sélection des allèles respectifs) sont rapportées. L'expression générale pour la fréquence des gènes en équilibre est:

$$\hat{q} = \frac{r + 2m - 1}{s + r}$$

Il est possible que des allèles tarés et d'autres, occupants des *loci* complexes, ayant une fréquence élevée dans des populations humaines puissent bénéficier d'un taux de ségrégation positif.

Zusammenfassung.

Die Untersuchung zeigt, daß rezessive, schädliche Allele (vier Letalgene und ein männliche Sterilität bedingendes Gen), welche in natürlichen Hausmauspopulationen vorgefunden wurden, eine große positive Spaltungsfrequenz aufweisen. Ungefähr 95 % der Spermien von heterozygoten Männchen ($+/t^w$) überführen das schädliche Allel. Der Verfasser veranschaulicht, wie solche Allele mit positiver Spaltungsfrequenz in einer Population bei Gleichgewichtsfrequenz trotz eines in entgegengesetzter Richtung wirkenden Selektionsdruckes beibehalten werden können. Diese Darstellung wird vom Verfasser in allgemeiner Form im Anhang wiedergegeben und die Gleichgewichtswerte für verschiedene Variationen von m (Spaltungsfrequenz), r und s (Selektionskoeffizienten für resp. Allele) werden angegeben. Der allgemeine Ausdruck für die Genfrequenz bei Gleichgewicht ist:

$$\hat{q} = \frac{r + 2m - 1}{s + r}$$

Der Gedanke liegt nahe, daß schädliche Allele und andere, im *complex loci* gelegene, welche eine hohe Frequenz in menschlichen Populationen haben, von einem solchen positiven Selektionswert für diesbezügliche Gameten begünstigt werden können.

LITERATURE

Chesley, P. and *L. C. Dunn:* Genetics *21*, 525, 1936. – *Dunn, L. C.:* Proc. Nat. Acad. Sci. *23*, 474, 1937. – *Dunn, L. C.* and *S. Gluecksohn-Waelsch:* Genetics (in press) 1953. – *Dunn, L. C.* and *W. C. Morgan, Jr.:* Proc. Nat. Acad. Sci. (in press) 1953. – *Gershenson, S.:* Genetics *13*, 488, 1928. – *Hiernaux, J.:* Annales du Musée Royal du Congo Belge, Tervuren (Belgique), Anthropologie *2*, 1, 1952. – *Morgan, T. H., C. B. Bridges* and *A. H. Sturtevant:* Bibliogr. Genetica *2*, 1, 1925. – *Neel, J.*, Blood (J. Hemat.), *6*, 389, 1951; *7*, 467, 1952. – *Silvestroni, E.* et al.: Nature, *165*, 682, 1950. – *Sturtevant, A. H.* and *Th. Dobzhansky:* Genetics *21*, 473, 1936. – *Wallace, B.:* Evolution *2*, 189, 1948.

APPENDIX

SOME EFFECTS OF VARIATIONS IN THE SEGREGATION RATIO AND OF SELECTION ON THE FREQUENCY OF ALLELES UNDER RANDOM MATING

By TIMOTHY PROUT

Denoting non-tailless alleles collectively by + and tailless alleles collectively by *t*, then let

q_n = the frequency of t-containing gametes in the gametic pool (frequency after segregation as opposed to frequency among the zygotes) of the generation n.

$1-q_n$ = the frequency of + containing gametes in the gametic pool of generation n.

$(1-r)$ = the average adaptive value of all male and female homozygotes of + alleles, +/+, relative to heterozygotes, +/*t*.

$(1-s)$ = the average adaptive value of all male and female homozygotes between t alleles, t/t, relative to heterozygotes.

m = average proportion of t-containing gametes among all gametes produced by +/t heterozygotes.

The arbitrarily assigned adaptive value of the heterozygotes, +/*t*, is taken as one.

s and r are the conventional selection coefficients which vary between zero and minus infinity. When the coefficient equals one, the designated zygote is lethal; when the coefficient equals zero, the zygote has adaptive value equal to the standard, in this case the heterozygote; when the coefficient is negative, the zygote has adaptive value superior to the heterozygote.

If there is random mating, the zygotes formed by these gametes, before the action of selection, will be in the following proportions, according to the terms of the expansion of $[q_n + (1-q_n)]^2$:

Zygote	+/+	+/*t*	*t*/*t*
Proportion	$(1-q_n)^2$	$2q_n(1-q_n)$	q_n^2

The ratio between the zygotes after the action of selection is:

Zygote	+/+	+/*t*	*t*/*t*
Ratio	$[(1-q_n)^2(1-r)]$:	$[2q_n(1-q_n)]$:	$[q_n^2(1-s)]$

In the gametic pool of these zygotes, that is, of generation n+1, taking into account the abnormal segregation, the ratio between the two resulting gametic classes is (showing the zygotic source of each term):

Zygotic Source $+/+$ $+/t$ t/t

Ratio $[(1-q_n)^2(1-r) + 2q_n(1-q_n)(1-m)] : [2q_n(1-q_n)m + (2q_n^2 1-s)]$

The proportion of gametes, in this gametic pool of generation $n+1$, which contain t alleles (denoted by q_{n+1}) is therefore,

$$q_{n+1} = \frac{q_n[2m + q_n(1-s-2m)]}{1-r+2q_n r-q_n^2(s+r)} \text{ and}$$

$$q_n - q_{n+1} = \triangle q_n = \frac{q_n[1-r+2q_n r-q_n^2(s+r)] - q_n[2m+q_n(1-s-2m)]}{1-r+2q_n r-q_n^2(s+r)}$$

At equilibrium, when $q_n = q_{n+1}$, let $\hat{q}$ denote the equilibrium proportion of t-containing gametes after segregation but before selection, then

when $\triangle q = 0$

$\hat{q} = 1$,

$\hat{q} = 0$, or

$-\hat{q}^2(s+r) - \hat{q}(2r-1+s+2m) + 1-2m-r = 0$

For the case which, as will be seen, is of the most practical interest, when the quantity $(s+r)$ is greater than zero,

when $(s+r) > 0$

$$\hat{q} = \frac{r+2m-1}{s+r}$$

Limiting the quantity $(s+r)$ to values greater than zero means that selection may be of the following types:

(1) One or the other alleles is a deleterious recessive (s or $r = 0$ and $0 < r$ or $s < 1$);

(2) One or the other alleles is deleterious in the homozygous condition but is also slightly semidominant. "Slightly" semidominant means here that the deleterious homozygote is more inferior to the heterozygote than the other homozygote is superior to the heterozygote ($0 < s$ or $r < 1$, r or $s < 0$ and $(s+r) > 0$);

(3) The two homozygotes are inferior to the heterozygotes ($0 < r$ and $s < 1$).

When selection is limited to the above types, then the following relationships between s, r, m, and $\hat{q}$ are true:

(1) In order that there be an equilibrium,

$$\frac{1-r}{2} < m < \frac{1+s}{2}$$

that is, if m is greater than the quantity $(1+s)/2$ then the effect of the favoring segregation will override the deleterious effects of

selection and cause the *t* allele to approach fixation. If m is less than the quantity (1—r)/2 then the effects of selection will override the effects of the favoring segregation and the t allele will be lost.

(2) If an equilibrium exists, under the conditions of (1), the equilibrium is stable.

(3) Regardless of the severity of selection there is always some value of m which can be found which will counteract selection producing an equilibrium. Also some value of m can always be found which can cause the t allele to approach fixation.

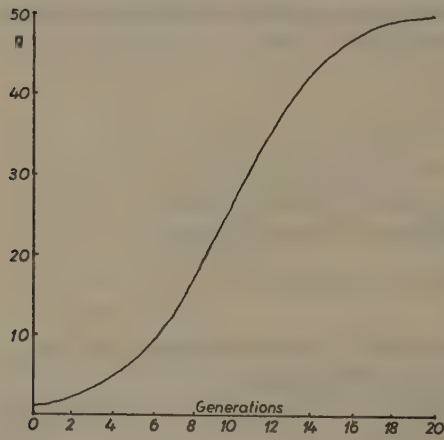


Fig. 1. The theoretical curve for the spread of a recessive lethal t allele (s = 1, r = 0) through a population. The initial proportion is q_0 = .01, and the segregation advantage is m = .75 which gives an equilibrium of $\hat{q}$ = .50. The expression for this curve is,

$$q_n = \frac{(2m)^n q_0}{1 + q_0\left[\frac{(2m)^n - 1}{2m - 1}\right]} \quad \text{when } m = \tfrac{1}{2}$$

$$q_n = \frac{q_0}{1 + nq_0} \quad \text{when } m = \tfrac{1}{2}$$

It appears that in the model used in arriving at the above conclusions considerations were not made for differences in m for the two sexes. Letting m denote the average segregation value between both sexes in the simplified model above yields identical results with a more complex model which considers the sexes separately. It should also be noted that variations in the sex ratio would not modify the results since, regardless of the sex ratio, each sex contributes exactly half the gametes for the next generation.

Table 1. The body of the table shows the equilibrium values of the proportion of t alleles after segregation but before selection, for different values of s and r, assuming m = .75 and (s+r) > 0.

					s					
0	.10	.20	.30	.40	.50	.60	.70	.80	.90	1.00	
	1.00	1.00	1.00	1.00	1.00	.8333	.7143	.6250	.5556	.5000	0
		1.00	1.00	1 00	1.00	.8000	.6667	.5714	.5000	.4444	—.10
			1.00	1.00	1.00	.7500	.6000	.5000	.4286	.3750	—.20
				1.00	1.00	.6667	.5000	.4000	.3333	.2857	—.30
					1.00	.5000	.3333	.2500	.2000	.1667	—.40 r
						0	0	0	0	0	—.50
							0	0	0	0	—.60
								0	0	0	—.70
									0	0	—.80
										0	—.90

The table above shows the theoretical equilibrium values for different values of s and r when m = .75. For equilibrium, s must be greater than .50, otherwise the t allele will approach fixation. Also r must be greater than —.50, otherwise the *t* allele will be lost.

A future paper is planned giving a more general treatment of the theoretical effects of abnormal segregation ratios in Mendelian populations. The effects of inbreeding will be discussed and also the cases where (s+r) is not restricted to values greater than zero.

Diamant, M.: Acta genet. *4*, 151–164, 1953

CHOLESTEATOMA
AN ANALYSIS OF FACT AND FIGMENT

By MARCUS DIAMANT, M.D. Halmstad, Sweden

Introduction.

A concept nowadays widely accepted is that cholesteatomata derive from different causes. Distinction is made between two groups of cholesteatomata. One group comprises all cholesteatomata localized in the aural region. These are said to originate from certain forms of otitis and are known as "secondary" cholesteatomata

which, in the modern view, are never neoplastic. The other group includes all those cholesteatomata which are situated elsewhere in the human body. They are regarded as neoplasms and are called "true" or "genuine" cholesteatomata. Both types develop into formations reportedly of identical appearance—macroscopically and microscopically. Thus the cholesteatoma is actually classified by its localization. And, as shown by the relevant literature, this classification is an arbitrary one. Our present knowledge of aetiology and development is most diffuse and incomplete. Particularly the belief that the aural cholesteatoma is a sequel to otitis seems derived from ill-founded hypotheses.

Actual problems.

The literature paints a captivating picture of the chronological evolution of today's *dualistic* attitude to the aetiology of the morphologically *uniform* cholesteatoma. The first encounter with the cholesteatoma was made at autopsy by pathologists. It had all the hallmarks associated with a neoplasm and was regarded as such. Hence the name cholesteatoma. Later on it was demonstrated by otologists that the aural region was by far the most common site of cholesteatoma. It was often found in so-called "chronic otitis", for the most part in cases with marginal perforation of the tympanic membrane, and at radical mastoidectomies. (Actually the presence of cholesteatoma is the most usual indication for such operations.) The fact that cholesteatoma so often is found together with marginal perforation of the tympanic membrane (which perforation is maintained to have been caused by "chronic otitis") is apparently the basis for the twin assumptions that an otitic process is the cause of the development of cholesteatoma, and that the cholesteatoma of the ear is consequently no neoplasm.

As a matter of fact it is now considered axiomatic that cholesteatomata of the ear without exception are due to otitic processes. Axiomatically this is said to preclude a neoplastic aetiology. Thus, had cholesteatoma occurred exclusively *within* the aural region, *present concepts* would have implied complete rejection of the alternative neoplastic aetiology. But cholesteatomata are in fact found at sites in the human body where it would be absurd to predicate an aetiology of otitis. This applies to those cholesteatomata which occasionally are found in the meninges, the orbit, the urinary tract, etc. That is why it has proved unavoidable *also* to retain the original view on a

neoplastic cholesteatoma. All *extra-aural* cholesteatomata are accordingly said to be neoplasms. Hence, had only *extra-aural* cholesteatomata existed, their being neoplasms or not would have been determined solely in accordance with the principles embodied in the current definition of a neoplasm. The conviction that an extra-aural cholesteatoma always constitutes a neoplasm has apparently had no repercussions on the presumption of a non-neoplastic aetiology of all cholesteatomata of the ear. However, were this presumption found inadequately based, unsubstantiated or even false, the neoplastic aetiology of cholesteatoma, wherever located, would not be proved thereby. The final verdict must always conform with current definitions of neoplasm. A thorough knowledge of the many problems concerned would, perhaps, make it still less justifiable to operate with two different modes of development of the morphologically uniform cholesteatoma. And that makes it worth while examining critically the foundation for the belief that cholesteatomata occurring in the aural region have an aetiology of otitis, as well as that according to which extra-aural cholesteatomata are neoplasms. In so doing, it is essential to find out whether it has been proved that cholesteatoma has an aetiology of otitis *at all*. If so, it must be decided whether cholesteatomata of the ear *always* have such an aetiology and, if so, whether this excludes a neoplastic aetiology. For this reason, a recapitulation of current views on neoplasm aetiology in general ought to be of value. Then it will be possible to consider the cholesteatoma in the light of those characteristics which are attributed to the neoplasm.

Conceptions of neoplasm aetiology.

As a rule the presence of a neoplasm is demonstrated macroscopically, both for benign and for malignant growths. To which of these types the neoplasm belongs is then decided after microscopical inspection. So far so good; but when we have to specify the reason for the development of a neoplasm we are manoeuvring wholly in the uncharted realm of conjecture. Numberless such hypotheses have been formulated down the years.

Virchow's [1863] irritation theory is based on the observation that growths can develop in response to constant, or sometimes occasional temporary, irritation. Such irritation may be chemical, thermal or mechanical. It may also take the form of chronic action by bacterial inflammatory processes or by factors associated with slowly healing

wounds with cicatricial changes, for example those developing after overdosage in roentgen therapy. The irritation is supposed to produce a local tendency to more rapid cell division at a rate far in excess of normal. The irritated cells are also said to assume more embryonic properties; and embryonic cells are capable of differentiation into a variety of tissues.

Cohnheim's growth theory rests on the hypothesis that during foetal life a single embryonic cell or cell group is removed from its proper environment and lodges at a foreign site in the body. Retaining their capacity to differentiate, these cells are then said to give rise to the subsequent development of growths.

Ribbert [1904] held a different view. He proposed that the abnormal rate of division and altered character of neoplastic cells was caused by inhibition of the normal growth restraining influence of the *surrounding* tissues.

The infection theory has also been advanced. It predicates the existence of a special cancerigenic virus, drawing the parallel with the proved existence of a transmissible virus capable of causing warts.

Cytochemical changes due to endocrine or exogenous factors are named as trigger mechanisms for the development of neoplasms. And modern biochemical as well as cellular tissue research generally has lately tried to get to the bottom of what actually happens in the cell when a neoplasm takes origin and develops. It is to be hoped that these efforts will be crowned with success and thus provide a more solid foundation for theories regarding the aetiology of neoplasms.

Heredity in general and local disposition of certain tissues have, too, been regarded as lying behind the evolution of certain tumours. Observations have been made which forbid an outright rejection of such factors.

In this connection *Bauer*'s [1949] mutation theory deserves mention. In his opinion neoplasms arise through mutations in the somatic cells. To this the objection can be raised that then neoplasms should occur more frequently in young persons, because their somatic cells divide more frequently than those of old persons.

One of the latest hypotheses was presented about ten years ago by *Dahlberg* [1940]. He likened the coming into existence of a neoplasm to the vegetative propagation of certain primitive organisms. Vegetative propagation is, of course, what happens when a cell or group of cells suddenly begins to divide more rapidly than other cells, resulting in the formation of a new individual. Only mature

organisms propagate themselves thus, however, and *Dahlberg* [1940] explained this by postulating that rapid cell division can occur only when the cells have already divided an adequately large number of times. Clearly, if *Dahlberg*'s [1940] hypothesis is correct, there is no mystery about the higher incidence of tumours in old people, nor about their development after a more or less intense Virchowian irritation such as compels a higher rate of cell division. Moreover, suggested *Dahlberg* [1940], vegetative propagation implies that a cell which thus begins rapidly to multiply simultaneously reverts to a more embryonic behaviour, acquiring the ability to differentiate to various kinds of tissues: a new individual thus develops from this one cell. In this light the embryonic type of neoplastic cells and, in certain cases, the development of variegated tissues, as in teratomata, take on a fresh perspective. Quite a few other phenomena associated with the appearance and development of tumours also seem capable of being brought into line with *Dahlberg's* [1940] theory of vegetative propagation.

Yet the large number of theories is in itself proof that these problems remain unsolved. The coming into existence and development of tumours is a sequence of events so numerous and disparate that no theory so far presented supplies all the pieces missing in our knowledge. Here, however, it is neither possible nor necessary to examine these questions in detail.

Otitis and Cholesteatoma.

When two disease processes are coexistent, the relations between cause and effect are often difficult to unravel. This problem is rendered even more complex by the fact that even in absence of all signs of concurrent otitis, aural cholesteatomata are said to have an aetiology of otitis. This is an automatic consequence of the aforementioned classification into (i) cholesteatomata of the ear by definition due to otitis and (ii) extra-aural cholesteatomata by definition neoplastic. This maxim is thus apparently unaffected by the fact that aural cholesteatomata are sometimes found without any signs of simultaneous otitis; nor by the fact that a *history* of otitis is lacking; nor by the fact that the ear in such cases can be proved *clinically and functionally intact*, apart from the cholesteatoma localized within the aural region. It is postulated in such cases that a hypothetical otitic process (i) could have occurred during intra-uterine life, or (ii) might have run a symptomless course, or (iii) perhaps presented symptoms

in such a slight degree that they went unnoticed. In all cases the otitis is supposed to have vanished leaving no trace except the primordium of a non-neoplastic cholesteatoma. Apparently the maxim is not even affected by the high incidence during childhood of otitic processes *with symptoms*. Such manifest otitides generally heal without giving rise to cholesteatomata. In certain individuals (those with comparatively small cellular systems) healing may be defective and a *central* tympanic membrane persists. *Marginal* perforation of the tympanic membrane—whose aetiology is wholly unknown according to the textbooks—has never been seen following childhood otitis of this type. *But marginal perforation of the tympanic membrane is generally held to be an irrevocable prerequisite to the development of an aural cholesteatoma.*

In what manner a cholesteatoma could arise and develop *following a clinically manifest otitis* is a problem in itself, and it has been much debated. The accepted view is (as presented by *Bezold* [1890], *Habermann* [1889] and *Politzer* [1891]), that the otitic process first (in an unspecified manner) gives rise to a marginal perforation of the tympanic membrane. The squamous epithelium of the external auditory meatus is thus provided with an entry by which it can invade the middle ear. Yet normal sloughing of the top layer of squamous epithelium continues. The desquamation products hypothetically cannot empty into the auditory meatus. Hence they are trapped and coalesce into a formation constituting the non-neoplastic cholesteatoma. This explanation contains a number of blatantly obscure points. Furthermore it eliminates in one stroke as causative factors both intra-uterine otitis and symptomless, healed otitis of early onset. For, according to this hypothesis, marginal perforation of the tympanic membrane must arise before cholesteatomata of the ear can be laid down. And marginal tympanic perforation have never been ascribed to intra-uterine so called foreign body otitis. Nor is their occurrence reported following healed otitis with a symptomless course in early life. It would moreover be a direct contradiction to claim that a symptomless otitic process has healed without leaving any traces, and that it has simultaneously produced a marginal perforation of the tympanic membrane. Actually no association whatsoever has been demonstrated between marginal perforation of the tympanic membrane and otitis. Another unexplained point is the formation of the preclosed matrix envelop containing the desquamation products. And what makes the

cholesteatoma grow? and by what mechanism? These points too are left in the dark.

Habermann's [1889] observation under the microscope of a direct transition from the squamous epithelium in the auditory meatus to that in the cholesteatoma matrix by way of the marginal tympanic membrane perforation constitutes a dubious proof of squamous epithelial growth *into* the middle ear *from* the auditory meatus. When there already is a bridge between the squamous epithelium in the auditory meatus and the matrix of the cholesteatoma, the direction of growth obviously cannot be determined by means of microscopical inspection. The reverse direction of growth could be claimed with equal justification. The sole condition would be the existence of squamous epithelium or of a performed cholesteatoma in the intact middle ear, before the development of a marginal perforation of the tympanic membrane commences. At the time of *Habermann*'s [1889] investigations such a possibility was not even considered. Subsequently, however, it has been shown not only theoretically possible but the true state of affairs in many cases.

Wittmaack [1918], having recourse to such publications, presented another explanation. He stated that squamous epithelium could enter the epitympanum via a plug-like infold from *Shrapnell*'s membrane (caused by negative pressure due to tubal occlusion). *Wittmaack* [1918] went on assuming that otitis sets in *after* these events and gives rise to a marginal perforation of the tympanic membrane. The growing cholesteatoma could then empty through the opening thus established. *Wittmaack*'s [1918] conception is clearly entirely different from *Habermann*'s [1889]. Yet it is in equal degree a construction unsupported by clinical findings and experiences. Both *Wittmaack* [1918] and *Habermann* [1889] have left unanswered the question of how the marginal perforation of the tympanic membrane comes into being. Marginal perforation of the tympanic membrane remains to be observed following otitis. Admittedly *Wittmaack* [1918] proposed that the infolding *as well as the necessary sequel of otitis* can take a symptomless course and perhaps occur during intra-uterine life. Apparently, according to *Wittmaack* [1918] the *prior existence* of a marginal perforation of the tympanic membrane is not *a sine qua non* for the development of the cholesteatoma as such. But for its emptying into the auditory meatus *Wittmaack* [1918] presupposes the existence of a perforation marginally in the tympanic membrane *which has to be caused by otitis.*

The causation of the marginal perforation of the tympanic membrane is however a problem which *Wittmaack* [1918] dealt with in passing only and really falls outside the scope of his conception of the aetiology of cholesteatoma. According to *Wittmaack* [1918] the aural cholesteatoma is not a neoplasm, nor has it strictly speaking an aetiology of otitis; it is rather a sort of cyst in the same category as posttraumatic inclusion cysts.

Neoplasm and Cholesteatoma.

In 1931 *McKenzie* published his long established conviction that all cholesteatomata are neoplasms (epidermoids) whose growth produces lesions on the inside of the tympanic membrane which, perforating it marginally, produce a fistula to the auditory meatus. Despite the factual argumentation and detailed analysis of *Politzer*'s [1891] interpretation of the clinical findings in aural cholesteatoma, and despite being seconded on a variety of points in the subsequent debate, *McKenzie*'s [1931] version seems to have received little attention. A large number of otologists have reported single or collected cases of cholesteatomata of the ear, stating that they must be primary, neoplastic cholesteatomata in view of the clinical picture and absence of all signs of otitis past or present. Actually these publications indirectly illustrate that the general view still is that aural cholesteatomata constitute non-neoplastic growths.

Marginal perforation of the tympanic membrane and Cholesteatoma.

In a series of paper published during the past fifteen years *Diamant* [1937] has insisted that the causation of marginal tympanic membrane perforation can be wholly separated from the discussion of whether or not the aural cholesteatoma is a neoplasm. In his view the possibility should be taken into account that the growing cholesteatoma, whatever its aetiology, is itself the factor giving origin to the marginal tympanic membrane perforation by inducing bone abrasion. In these respects, then, the cholesteatoma has properties corresponding to those presented by all expansive tumourous processes and inclusion cysts. For the present discussion it matters little whether the bone damage is due to pressure or to chemical action. *Diamant* [1948], and before him *McKenzie* [1931], have drawn attention to the similarity between the causation of a marginal

tympanic membrane perforation and the causation of a labyrinthine fistula, a fistula to the facial canal or an exposure of the dura following destruction of the internal cortical wall of the temporal bone in cases of large occlusive cholesteatomata. In growing cholesteatoma there is, according to *Diamant* [1948], no reason for playing with two different modes of bone destruction in the auditory region. Not least is this unreasonable because marginal tympanic membrane perforation is unknown after otitis, and so is a central perforation that has progressed to a marginal one. It must be more far-fetched to attribute the marginal perforation of the tympanic membrane to a purely conjectural otitis without signs and symptoms, past or present, than to admit the analogy between these various bone abrasions, since abrasion of bone is the known cause of at any rate labyrinthine and facial canal fistulation and dural exposure in cases of cholesteatoma.

Actually *Diamant*'s [1948] conclusions from personal clinical observations regarding the genesis of marginal perforations are not in themselves imcompatible with *Wittmaack*'s [1918] hypothetical invagination of *Shrapnell*'s membrane. But for this particular problem it matters not in the least *how* the primordial cholesteatoma is laid down. *Habermann*'s [1889] theory, however, is based on the assumption that marginal perforation of the tympanic membrane *precedes* the establishment of the cholesteatoma. Such an assumption lacks clinical foundation. Other hypotheses relating to the origin of the primordial cholesteatoma seem independent of a preexisting marginal perforation. *Wittmaack*'s [1918] and *McKenzie*'s [1931] theories excepted, this particular problem has been disregarded. Nevertheless the view is generally held that otitis somehow brings about the marginal perforation of the tympanic membrane. On this point, therefore, *Habermann*'s [1889] and also *Wittmaack*'s [1918] conjectures have been generally accepted. Only *McKenzie* [1931] and *Diamant* [1948] clearly state that otitis has no place in the aetiology of marginal perforation of the tympanic membrane.

Squamous epithelium and Middle ear.

Obviously the ultimate aim of both *Habermann* [1889] and *Wittmaack* [1918] was to present a hypothesis that would plausibly explain the presence of squamous epithelium in the middle ear. *Habermann* [1889] held that squamous epithelium proliferates inwards from the auditory meatus (through a perforation preformed in

the tympanic membrane); *Wittmaack* [1918] maintained that squamous epithelium is drawn into the middle ear, not directly from the auditory meatus but from the outer surface of the tympanic membrane facing the auditory meatus. But these are mere differences of detail, the same fundamental basis regarding the "source" of the squamous epithelium which later gets into the middle ear underlies both *Habermann*'s [1889] and *Wittmaack*'s [1918] theories. They diverge only as regards the necessity of a pre-existing perforation of the tympanic membrane. The marginal tympanic perforation is a *conditio sine qua non* for *Habermann* [1889]; not so for *Wittmaack* [1918]. However, they have the same attitude to the causation of the marginal perforation, namely that it is a sequel to otitis. These differences and similarities are often mixed up, so that *Wittmaack*'s [1918] theory is accepted and at the same time *Habermann*'s [1889] concepts are thought applicable to it. Meanwhile it is forgotten that neither has shown by clinical findings and experiences that otitis ever gives rise to marginal perforation of the tympanic membrane.

Even more hypotheses have been put forward in explanation of the presence of squamous epithelium in the middle ear, however. Thus it has been stated that metaplastic production of squamous epithelium can take place in response to irritation by a chronic otitic condition. *Kelemen* [1934] suggested a traumatic aetiology, particularly parturition traumatism. *Teed* [1936] has published investigations indicating that cell islets of squamous epithelium are a "normal" occurrence in the middle ear. He assumed that they were associated with the embryonic development of the ear from the ectoderm and early gill clefts.

Admitting the existence of squamous epithelial cells in the middle ear, one can proceed along alternative lines. Proliferation of these cells may give rise to a neoplastic or a non-neoplastic cholesteatoma, in the former as well as in the latter event by the same means as choleastomata are developed in, say, the pia mater, or in other tissues where tumours of ectodermal origin occur, which are maintained to be neoplastic.

Bergstrand [1943] emphasized the similarity of those cell proliferations that are associated with inflammatory disturbances and those that characterize a neoplasm. Hence it is difficult to draw the dividing line between hyperplasia and neoplasm. He stated that the difference is that inflammatory proliferation stops growing after the end of the inflammatory process, whereas the neoplasm goes on

growing whether or not the precipitating factor is still present. Taking this difference for granted, however, one will find it very difficult if not impossible to assert that the cholesteatoma of the ear is a growing inflammatory hyperplasia, although there is not a sign of active otitic infection. The expansive growth of the cholesteatoma does, on the other hand, conform exactly to that of the neoplasm. (Even so *Bergstrand* [1943] seems to have thought *Habermann*'s [1889] otitis-invasion hypothesis the sole acceptable explanation of the development of aural cholesteatomata. At any rate he made no intimation to the contrary.)

The neoplastic nature cannot be drawn in doubt by the fact that an otitic condition may be necessary *to set going* the growing process as such. It has been shown that neoplastic growth sometimes can commence as a result of irritation by extrinsic factors. The neoplastic aetiology is not even invalid in the event that the primordium owes its *origin* to external irritants. The latter mode of development is postulated by those who maintain that the squamous epithelium is produced through a metaplastic response to the irritation. Even primordia of certain neoplasms are laid down in response to irritation, e.g. cancer of the lips in pipe smokers. Thus, if the irritating factor assumedly constituting the extrinsic aetiology of the neoplasm should prove to be otitis (and why should it be fundamentally wrong to suppose that chronic irritation owing to an inflammatory otitic condition may constitute and represent this factor), then it cannot be contended that the cholesteatoma of the ear is not neoplastic merely on the grounds that it has developed through the action of otitic processes. (This does not mean that the otitic irritation has actually been shown to exist.)

Nor can it be claimed that the aural cholesteatoma has an aetiology other than that of the extra-aural cholesteatoma merely because the former is so much more common. Indeed all tumours occur with varying frequency at different sites. And the cholesteatoma does not differ so overwhelmingly in this respect from other tumours that this could not be due to the very complexity of the embryonic development of the ear and to the fact that aural cholesteatomata are probably recorded more frequently than cholesteatomata elsewhere.

Not even a concomitant otitic process can simply be named the precipitating "extrinsic" factor. Such otitis could either be con-

current accidentally or it might be a secondary effect of the presence of a neoplastic cholesteatoma.

The not infrequent bilateral presence of cholesteatomata has been advanced as an argument against the neoplastic theory. There is no foundation for such a view. Often the aetiologically obscure exostoses in the auditory meatus are also bilateral. The same applies to congenital deformities of paired organs, e.g. cysts, lateral fistulae of the neck, supernumerary mamillae, etc. In so far as one accepts *Cohnheim*'s old tumour theory as a basis for the laying down of primordial neoplasms during fetal life, this theory combined with congenital deformities seems capable of giving origin to cholesteatomata. Both the bilaterality and frequency distribution of cholesteatomata would thereby receive an explanation.

Finally, even were it possible to state definitely that an otitic process can give rise to a primordial cholesteatoma, indeed even if it were conclusively demonstrated that *Habermann*'s [1889] inward growth actually gives rise to a cholesteatoma in the manner specified, it would by no means be proved that *all* aural cholesteatomata have an aetiology of otitis. One must of course also take into consideration that the cholesteatoma, if it can occur as a true neoplasm to one site, must be capable of existing as a true neoplasm in the aural region too. Thus, one must expect some of them to be such neoplasms as occur outside the region of the ear.

On the other hand the neoplastic nature of cholesteatomata has not been proved. But the same applies to most forms of neoplasms, e.g. cancer, where the virus' aetiology remains on the list of undiscarded theories. The cholesteatoma does, however, in all respects conform to current definitions of a neoplasm, which applies to cholesteatomata of any location. Consequently there is no justification for distinguishing between cholesteatomata of the ear and other cholesteatomata, at least not in respect of any assumed differences in aetiology, or as regards its neoplastic or non-neoplastic nature.

The position can be stated thus: In conformity with accepted views one can hold that on the one hand "chronic otitis" constitutes the aetiology of *non-neoplastic* cholesteatomata in all cases of cholesteatoma of the ear. Perhaps a finer distinction could be drawn by confining the otitic aetiology to clinically manifest otitides. In either case, however, disparity arises on several points between the clinical findings and the theoretical considerations pertaining to

such disease processes. On the other hand one could agree with *Virchow* [1863] and *McKenzie* [1931] in maintaining that the neoplastic aetiology is common to all cholesteatomata irrespective of localization and the existence of otitis. By so doing one at any rate avoids introducing inconsistencies between clinical findings and current definitions of neoplasms. But this does not prove that the cholesteatoma ever or always is a neoplasm. As a rule it is difficult to prove medical phenomena. Perhaps the choice of alternative is a question of temperament more than anything else. Nevertheless our choice—if a choice is at all necessary—should be based on our own interpretation of available facts rather than slavishly concur with those of text-book interpretations and current classifications which do not present the whole picture for our study.

Summary.

To-day, cholesteatoma is mainly classified into two groups, primary and secondary cholesteatoma. "Primary" cholesteatoma means neoplasm. "Secondary" cholesteatoma means localization to the middle ear spaces. Thus, the basis for the classification is neither uniform, nor logical. The maintenance of an otitis, not to say a marginal perforation of the tympanic membrane being a pre-requisite for the development of the onset (Anlage) of aural cholesteatoma is just as unproved as is the hypothesis that they all are non-neoplasms or neoplasms. Our lack of knowledge as to whether the aural cholesteatoma is a neoplasm or not, should not, however, obscure our views concerning the development of the marginal perforation of the tympanic membrane. The aural cholesteatoma (uniform or not at onset) may always be prior ("primary") to the marginal perforation of the tympanic membrane. The cholesteatoma, if located to the middle ear spaces and their neighbourhood may (in ears with a small air cell system) cause the marginal perforation by its own growth. The presence of a marginal perforation can, therefore, not be used as a support for the "secondary" cholesteatoma hypothesis.

Résumé.

A présent, les cholestéatomes auriculaires se rangent principalement en deux groupes, les cholestéatomes «primaires» qui sont des néoplasmes, et les cholestéatomes «secondaires», ce qui revient à dire que la tumeur se trouve dans l'oreille moyenne. Ainsi cette division manque aussi bien d'unité que de logique. Que la présence d'une otite ou bien d'une perforation marginale de la membrane du tympan soit une condition essentielle pour le développement d'un cholestéatome auriculaire est un hypothèse entièrement gratuite. Parfois le cholestéatome auriculaire peut se développer avant (être «primaire») la perforation marginale de la membrane du tympan. Si le cholestéatome se trouve dans l'oreille moyenne ou dans une partie voisine, il peut, surtout dans les oreilles ayant un système cellulaire maigre, provoquer une perforation marginale lorsqu'il se développe. Ainsi, une perforation marginale du tympan existante ne peut pas supporter l'hypothèse des cholestéatomes «secondaires».

Zusammenfassung.

Cholesteatom wird zurzeit hauptsächlich in zwei Gruppen aufgeteilt, nämlich «primäres» Cholesteatom, womit man Neoplasmus angibt und «sekundäres» Cholesteatom, welches bedeutet, daß das Cholesteatom im Mittelohr lokalisiert ist. Diese Aufteilung ist daher weder einheitlich noch logisch. Es ist eine vollkommen unbewiesene Hypothese, daß das Vorkommen von Otitis, um nicht zu sagen marginaler Perforation des Trommelfelles, eine Voraussetzung für ein beginnendes Cholesteatom im Ohr sei. Cholesteatom im Ohr kann zuweilen früher auftreten («primär» sein) als die marginale Perforation des Trommelfelles. Wenn das Cholesteatom im Mittelohr oder den angrenzenden Gebieten lokalisiert ist, kann es durch sein Wachstum bei Ohren mit kleinem Zellsystem eine marginale Perforation verursachen. Das Vorkommen von marginaler Perforation kann deshalb als Stütze für die Hypothese von «sekundären» Cholesteatomen nicht verwendet werden.

REFERENCES

Bauer, K. H.: Das Krebsproblem. Springer, Berlin/Göttingen/Heidelberg 1949. – *Bergstrand, H.:* Lärobok i patologi, 1943. – *Bezold, F.:* Z. Ohrenheilk. *20*, 1890. – *Dahlberg, G.:* Uppsala LäkFören Förh. N. V. *46*, 21, 1940. – *Diamant, M.:* Acta oto-laryng., Stockh. *25*, 507, 1937; Arch. Otolaryng., Chicago *39*, 581, 1948; «Chronic Otitis». S. Karger, Basel 1952. – *Habermann, J.:* Arch. Ohr.-, Nas.- Kehlk.heilk. *27*, 1889. – *Kelemen, G.:* Acta oto-laryng., Stockh. *20*, 211, 1934. – *McKenzie:* Proc. Roy. Soc. Med. *24*, 332, 1931. – *Politzer, A.:* Wien. med. Wschr. 1891. – *Ribbert, H.:* Geschwulstlehre, 1904. – *Teed, R. W.:* Arch. Otolaryng., Chicago *24*, 455, 1936. – *Virchow, R.:* Geschwülste, I/II, 1863. – *Wittmaack, K.:* Über die normale und pathologische Pneumatisation des Schläfenbeines. Jena 1918.

Aus dem Institut für gerichtliche und soziale Medizin an der Universität Kiel
(Direktor: Prof. Dr. med. W. Hallermann)
und dem Max-Planck-Ins.itut für vergleichende Erbbiologie und Erbpathologie,
Berlin-Dahlem
(Direktor: Prof. Dr. phil. H. Nachtsheim)

DER MAKAK-TYPUS IM TASTLEISTENSYSTEM EINER DEUTSCHEN SIPPE

Von G. GEIPEL und W. LEHMANN

Gelegentlich einer erbbiologischen Untersuchung in einem Fall, in dem die Ehelichkeit eines Kindes angefochten wurde, erhob *W. Lehmann* bei der Untersuchung der Tastleisten der Hände und Füße einen Befund auf den Händen des Kindes, der wegen seiner Besonderheit und großen Seltenheit eine Veröffentlichung rechtfertigt.

Die *Probandin*, Ellen D. (Nr. IV, 1 der Sippentafel, Abb. 1), die bei der Untersuchung 3,6 Jahre alt war, ist das einzige Kind des

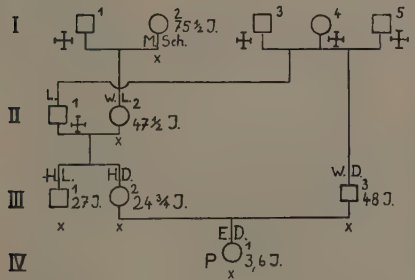


Abb. 1. Stammtafel der Sippe. x = selbstuntersucht. P = Probandin.

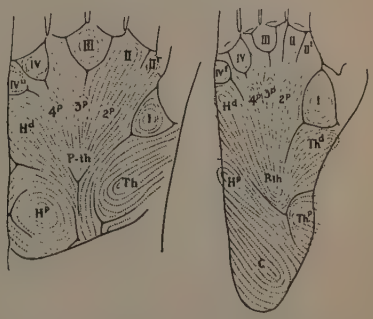


Abb. 2. Grundplan für Hand und Fuß der Primaten (nach *Midlo* und *Cummins*). Der Plan zeigt die Ballenanordnung, interdigital I-IV^{u}, Thenar Th, Hypothenar distal H^{d}, prox. H^{p} und Parathenar P-th 2^{p}-4^{p} für prox. offene Schleifen.

Ehepaares D. Die Geburt erfolgte rechtzeitig; es waren alle Zeichen der Reife vorhanden. Seit einigen Monaten treten bei dem Kinde im Abstand von etwa 4 Wochen Anfälle auf, die nach der Schilderung der Großmutter des Kindes mit plötzlichem Hinstürzen beginnen und etwa eine Minute anhalten, wobei Zuckungen am ganzen Körper auftreten. Danach verfällt das Kind in einen Schlafzustand. Gelegentlich soll es auch zum Einnässen und Schaumbildung am Mund kommen. Es dürfte sich um echte epileptische Anfälle handeln.

Die körperliche Untersuchung des Kindes bot sonst keine besonderen Befunde, insbesondere fanden sich an Rumpf oder Gliedmaßen keinerlei Mißbildungen oder sonstige Auffälligkeiten. Es besteht lediglich ein Einwärtsschielen, wie man das bei Kleinkindern nicht selten beobachtet.

Die *Mutter der Probandin*, H. D., 24¾ Jahre alt (Nr. III, 2), leidet seit ihrem 3. Lebensjahr gleichfalls an offenbar epileptischen Anfällen, die aber jetzt seltener auftreten. An Händen und Füßen bestehen keinerlei Mißbildungen. Psychisch bot sie keine Auffälligkeiten.

Der *Vater der Probandin*, W. D. 48 Jahre alt (Nr. III, 3), übrigens der Halbbruder von II, 1 und daher der Halbonkel seiner Ehefrau, ist gesund und zeigt an den Gliedmaßen keinen von der Norm abweichenden Befund. Bei der erbbiologischen Ähnlichkeitsuntersuchung konnte er als Erzeuger des Kindes (er hatte die Ehelichkeit des Kindes angefochten) nicht ausgeschlossen werden.

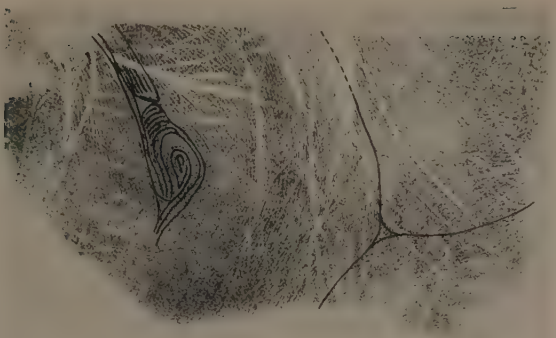
Von einigen weiteren Angehörigen aus der Familie der Mutter der Probandin konnten z. T. Abdrücke der Hände und Füße gewonnen werden.

Der *Onkel der Probandin*, Bruder der Mutter, H. L., 27 Jahre alt (Nr. III, 1), ist gesund. Aus äußeren Gründen konnte nur eine Inspektion der Hände vorgenommen werden. Den bei der Probandin beschriebenen Wirbel im III. Interdigitalraum weist er nicht auf. Er besitzt in den Interdigitalräumen II-IV beiderseits Schleifen, wobei die Schleifen im IV. Interdigitalraum im Kern eine Einrollungstendenz aufweisen und wie im II. mit Nebentriradien versehen sind.

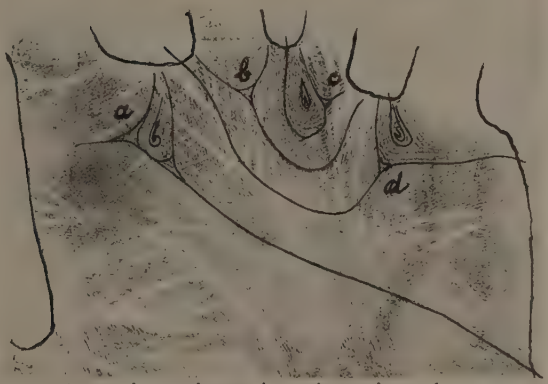
Die *Großmutter der Probandin*, W. L., 47½ Jahre alt (Nr. II, 2), ist gesund. Eine körperliche Untersuchung ließ sich nicht durchführen. An den Gliedmaßen fanden sich keine Mißbildungen.

Ihr Ehemann (Nr. II, 1), ebenso wie dessen Eltern (Nr. I, 3 und I, 4) leben nicht mehr.

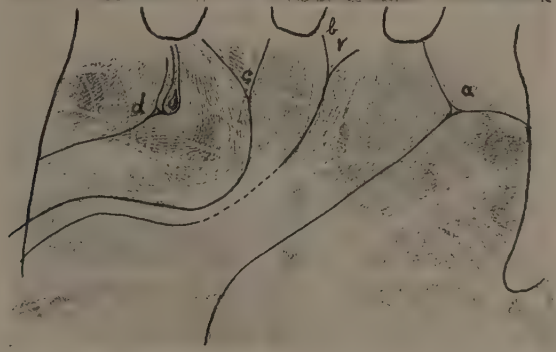
Schließlich konnten noch Handabdrücke bei der *Urgroßmutter*



3. *Bettmann*-
Thenarmuster
er rechten Hand
der Urgroßmut-
Nr. 4).



4. Drei distale
digi alwirbel
er rechten Hand
Urgroßmutter
4). Makaktypus.



5. Nur *ein* In-
gitalwirbel in
distalen Region
Urgroßmutter
nifestationshem-
g) links.

der Probandin, M. Sch. 75 ½ Jahre alt (Nr. II, 2), die noch sehr rüstig ist, gewonnen werden. An den Händen hatte sie keine Deformitäten.

Ihr Ehemann (Nr. I, 1) lebt nicht mehr.

Zum *Tastleistenbefund* auf den vorliegenden Abdrücken äußert sich *G. Geipel* auf Anregung durch *W. Lehmann* folgendermaßen:

1953, dem Jahre dieser Veröffentlichung, werden es gerade 70 Jahre, daß Anatomen, Anthropologen und Zoologen sich mit der Erforschung der Tastballen und Tastleisten bei Tieren und Menschen eingehender beschäftigt haben. Seitdem sind nach einer von *Midlo* und *Cummins* angegebenen Bibliographie von rund 50 Verfassern etwa 100 der wichtigsten Arbeiten darüber veröffentlicht worden, angefangen von dem Deutschen *A. Kollmann*, der 1883 das Buch «Der Tastapparat der Hand der menschlichen Rassen und der Affen» herausbrachte, bis zu den US-Amerikanern *H. Cummins* und *Ch. Midlo*, deren Buch «Finger Prints, Palms and Soles» 1943 erschienen ist. Als grundlegendes morphologisches Ergebnis aller Vorarbeiten konnten die beiden letzteren 1942 in ihrem Buche «Palmar and Plantar Dermatoglyphics in Primates» einen Grundplan von Handfläche und Fußsohle aufstellen, der in Abb. 2 wiedergegeben ist. Auf ihn und eine Reihe von äffischen Handabdrücken, die in engster Anlehnung an die Originale zeichnerisch festgehalten worden sind, sollen die bei der Sippe aufgefundenen Merkmale bezogen werden. Die Bilder der Affenhände sind unter Vergrößerung dem Buche von *Midlo* und *Cummins* entnommen. Alle Hände erscheinen wie in dem Buche als rechte Hände oder Abdrücke linker.

Urgroßmutter (Nr. 4)

Handformel: Rechts: 11–7.9.7.5'–t'–O/A^{c}.W/q/1$^{r}_{v}$.W.W.W$_{v}$

Links: 7.5".5'.3–t'–O/A^{c}.W/q/L^{r}.m.O.W$_{v}$

Die Muster W/q/1^{r} bzw. W/q/L^{r} liegen auf den mit Th und I bezeichneten Tastballen des Grundplans (Abb. 2). Außer diesen Wirbeln aber, und das ist besonders auffällig, liegen rechts 3 Wirbel in den Interdigitalen II, III und IV und links *ein* Wirbel in J IV. Während die Thenarwirbel meist gut manifestiert sind (Abb. 3), ist die Expression der interdigitalen Wirbel in der distalen Region nicht immer vollkommen (Abb. 4 und 5). Oft ist nur eine einzige Spiralwindung zu erkennen, doch muß sie als hinreichender Unterschied gegen eine distale Schleife angesehen werden. Der Grund hierfür ist im Raummangel der kleineren distalen Ballen zu sehen. In Anbetracht des hohen Alters der Urgroßmutter sind die Handleisten durch grobe

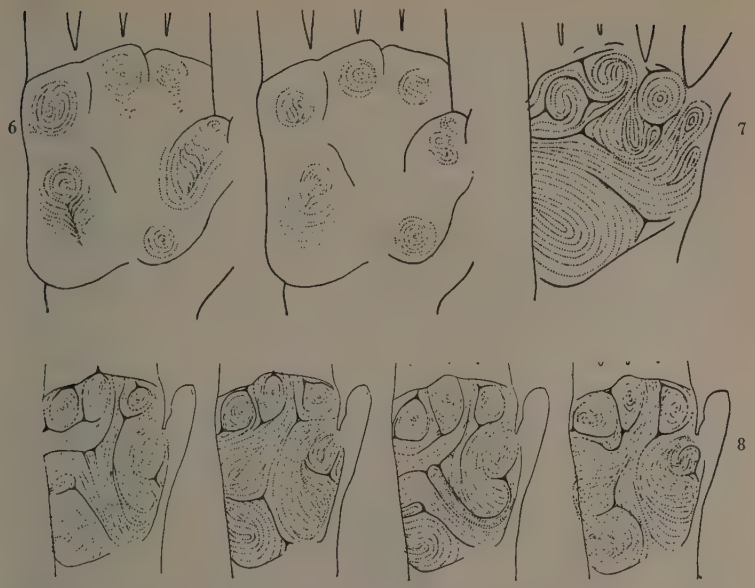
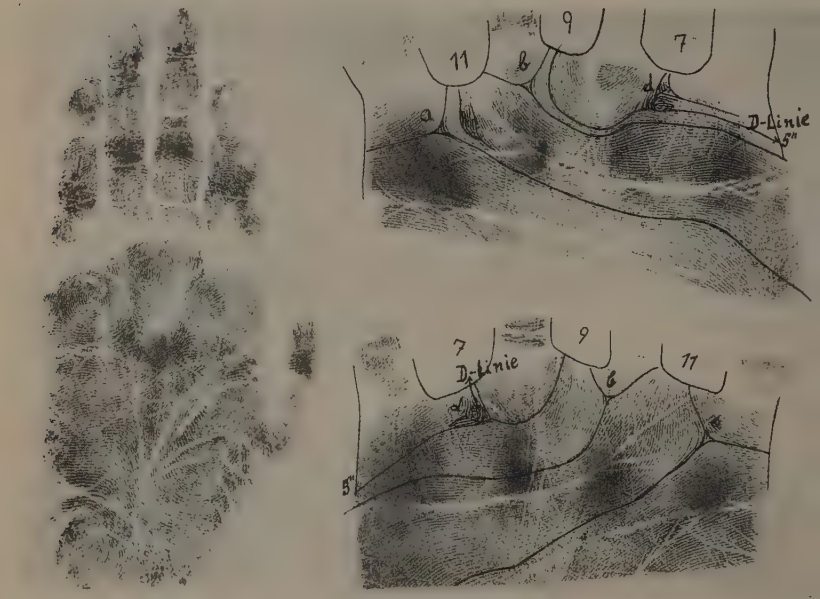


Abb. 6. Zwei etwas von einander verschiedene Galago-Hände mit distalen und proximalen Wirbeln. – *Abb. 7.* Brüllaffen-Hand mit disatlen Wirbeln und Hypothenarschleife. – *Makaktypus. Abb. 8.* Vier Makak-Hände, die alle die typischen drei distalen Interdigitalwirbel zeigen. Im Interdigitalraum I haben der erste und vierte Wirbel, der dritte und vierte Schleifen. Der dritte hat eine Parathenarschleife in 4^{P}. Thenar und Hypothenar tragen zumeist Schleifen.

Arbeit zerstückelt und mußten im Abdruck durch getreue Nachzeichnung etwas ergänzt werden, um im Druck erkennbar zu sein[1]. Drei interdigitale Wirbel, distal gelegen, sind beim Menschen eine so große Seltenheit, daß eine Nachforschung geboten erschien, wo sich Gleiches oder Ähnliches in der Primatenreihe finden ließe. Es ergab sich, daß man bis zu den Prosimiern hinabsteigen muß, um etwa bei *Galago* (Abb. 6) *zuluensis* drei distale Wirbel zu finden. Dabei soll von wenigen tiefer stehenden Nicht-Primaten abgesehen werden. In der Simierreihe sind sie bei Alouatta (Brüllaffe, Abb. 7) in Abb. 81 ff.

[1] Die Einzeichnungen von Tastleisten in die Abdrücke sind unter Anleitung des Verfassers von dem Biologiestudenten *Günther Gerisch* getreu der Vorlage ausgeführt worden. Die Fotos hat *Maria Mikolajczyk* hergestellt.

der genannten Quelle zu finden. Dann folgen Saimiri(Totenköpfchen), Aotus (Nachtaffe), Cebus (Kapuziner), Papio (Hundsaffe), *Pithecus* (*Makak*, Abb. 8 und 9) und schließlich Pygothrix (Langur). Weiter aufwärts werden diese Wirbel immer seltener, namentlich in J III, so daß ihr Auftreten beim Menschen mit Recht als *Makaktypus* bezeichnet werden darf. *H. Wilder* hat als erster im Biological Bulletin auf die Makakhand als einer Ausgangsform für die stammesgeschichtliche Ableitung auch der Menschenhand hingewiesen. Eine Wiederholung dessen findet sich in dem Buche «Personal Identification» von *Wilder* und *Wentworth*, illustriert durch zeichnerische Betonung der wesentlichsten Handleisten, S. 122, Abb. 36.



9

Abb. 9. Abdruck einer linken Hand von Makak, die in J II den Übergang des Wirbels zur Schleife mit proximaler Öffnung zeigt. – *Abb. 10.* Abdruck der rechten Hand der Großmutter mit einer D-Schleife in IV", die sich nach ulnar öffnet, so daß die D-Linie nach 5" ausläuft. Die Nebenlinie zieht nach 9. Die C-Linie fehlt. Die distalen Ballen sind hier flach. – *Abb. 11.* Abdruck der linken Hand der Großmutter mit einem einzigen Wirbel in IV. Die D-Linie läuft distal aus nach 7, ihre Nebenlinie nach 9. Die C-Linie fehlt. Die distalen Ballen sind hier flach.

Großmutter (Nr. 6).

Handformeln. Rechts: 9–5".0.5".5'–t'–O/A^{c}.O.D$_{v}$. O.D^{u}
Links: 9–7 0 5".5'–t'–O/A^{c}.O.m.O.W$_{v}$

Die Formeln besagen, daß beiderseitig Muster mit Nebentriradien im IV. Interdigital liegen, von denen das auf der rechten Hand (Abb. 10) eine sehr seltene, nach *ulnar* offene Schleife D^{u}, das auf der linken Hand (Abb. 11) ein nach distal offener, allerdings nicht vollkommen manifestierter Wirbel W ist. Die Schleife D^{u} liegt nicht, wie es nach der Formel scheint, im distalen J IV, sondern auf dem Ballen IV^{u} des Grundplans. Daher läuft die Hauptlinie D, die in d entspringt, nicht nach distal, sondern nach 5" am *ulnaren* Handrande aus. Aus einer vom Verfasser aufgestellten Liste von Endigungstypen (Modaltypen), die sich auf rund 4300 Hände von 16 Rassen erstrecken, ist ein solcher ulnarer Auslauf nur ein einziges Mal vermerkt, und zwar bei einem Chinesen, ein Beweis für die Seltenheit. Die in den Nebentriradien entspringenden Nebenlinien ziehen nach 9.

Mutter (Nr. 8).

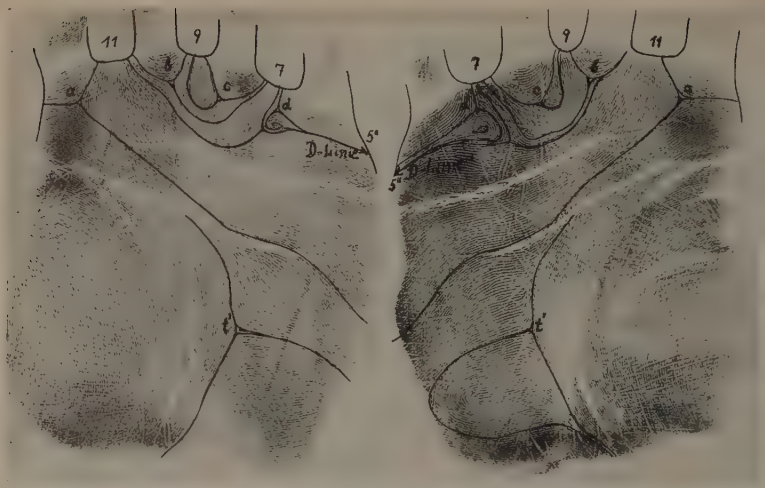
Handformeln Rechts: 11–5".9.7.5'–t'–O/A^{c}.O.O.L.W.
Links: 11–5".9.7.5'–t'–O/L^{c}.O.m.L.W

Die Hauptlinie D läuft auf beiden Händen wie bei ihrer Mutter (Nr. 6) auch nach 5", während die von den Nebentriradien ausgehenden Nebenlinien nach 11 ziehen (Abb. 12 und 13). Die beiden Wirbel W sind wie die Schleife D von Nr. 6 nach ulnar offen; sie liegen wie jene Schleife auf dem Ballen IV^{u} des Grundplans. Wegen der großen Seltenheit solcher Lage beim Menschen kann hier kein Zufall vorliegen. Man muß Erblichkeit für wahrscheinlicher halten.

Vater (Nr. 9).

Handformeln. Rechts: 9.9.5".4–t–L$^{r}_{v}$.O.O.L.O
Links: 9.9.5'.3–t–L$^{r}_{v}$.O.O.L.M

Hauptlinien und Muster bieten nichts Auffälliges, doch soll beiläufig erwähnt werden, daß die Dreifingerfurche auf beiden Händen sich an ihrem Ende in einen distalen und einen radialen Ast gabelt, und daß der radiale zur Fünffingerfurche strebt, also eine Brücke bildet, so daß deutliche Tendenz zur Bildung einer *Vierfingerfurche* besteht.



12

Abb. 12. Abdruck der rechten Hand der Mutter (Nr. 8), der einen Wirbel in IV^{u} mit Auslauf der D-Linie nach 5" zeigt. Der Thenar ist musterfrei, ebenso J I, wie bei Nr. 6. – *Abb. 13.* Abdruck der linken Hand der Mutter (Nr. 8), der auch einen nach 5" mit der D-Linie auslaufenden Wirbel in IV^{u} zeigt. Der Thenar ist, wie auch J I, musterfrei wie bei Nr. 6.

Urenkelin, die Probandin (Nr. 10).

Handformeln. Rechts: 11–11.9.8.5'–t'–O/L^{c}.L^{c}/q/L^{c}.O.W.W
Links: 11.(9).7.3^{h}–t'–W/A^{r}.L^{c}/Q/L^{r} .O.W.W

Auf beiden Händen Abb. 14 und 15 sind die Thenare wie bei der Urgroßmutter mit dem dreiteiligen *Bettmann*schen Muster besetzt und decken die Ballen Th und I des Grundplans. Die Interdigitale III und IV tragen überraschend große, mit *drei* Triradien ausgebildete Wirbel, von denen der in J III rechts von der sich deckenden B- und C-Linie umschlossen wird, während der in J III links die C-Linie (in der Formel mit [9] bezeichnet) in seinen nahezu kreisförmigen Wirbelstrom aufnimmt; es kommt hierbei zu keinem Auslauf am Handrande, was sehr selten sein dürfte. Der seltene dritte Triradius liegt *zwischen* den Haupttriradien b und c, bzw. c und d. Die zentrale Lage dieser beiden Wirbel auf J III ist in ihrer Formschönheit so auffallend und selten, daß sie nur mit der bei den genannten Halbaffen und Affen verglichen werden kann. Bei einzelnen Makak-

14 15

Abb. 14. Rechter Handabdruck der Urenkelin (Nr. 10) mit den zwei Wirbeln in J II und J III; der dritte, seltene Triradius liegt an der Schwimmhaut. Er kommt auf der Menschenhand fast nie vor. Das Thenarmuster ist wie bei Nr. 4 vorhanden. – *Abb. 15.* Linker Handabdruck der Urenkelin. Es sind die entsprechenden Muster wie rechts vorhanden; auffallende Symmetrie. Nur die Hypothenare weichen in der Lage der Schleifen etwas voneinander ab. Die Übereinstimmung mit der Urgroßmutter, Überspringen zweier Generationen, manifestiert sich selten so vollkommen.

exemplaren (Abb. 16) sind die Wirbel in J II durch Schleifen ersetzt; aber es kommen da auch offene Felder vor. Die Manifestation des Typus kann also unvollständig sein. Bei unserer Probandin gilt das für J II, wo ein offenes Feld O vorliegt. Bei ihr zeigt sich der Manifestationsmangel in der *Anzahl* der Wirbel, bei der Urgroßmutter in der *Expression*. Dies ist um so weniger zu verwundern, als bei den höher

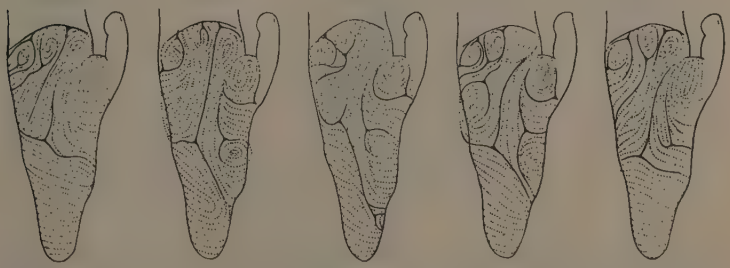


Abb. 16. Fünf verschiedene Variationen des Makaktypus auf den Hinterhänden: Häufige Schleifen in J II, Schleifen und Wirbel in J I, nach ulnar offene Wirbel in J IV auf dem ersten und dritten Fuß.

stehenden Gibbons und den Menschenaffen nur noch gelegentlich, d.h. nur bei einzelnen Exemplaren, also *un*typisch, distale Wirbel auftreten, z.B. bei Pongo (Orang Utan, Abb. 17). Ähnlich bei Gorilla (Abb. 18) und bei Pan (Schimpanse, Abb. 19). Es verdient hervorgehoben zu werden, daß bei Pan die Schleife in J II einen Nebentriradius hat wie stets beim Menschen, falls er sie besitzt. Einfache Schleifen in J II kommen bei keiner Menschenrasse vor.

Da der Makak*typus* in unserer Sippe sich nur bei der Urgroßmutter und ihrer Urenkelin manifestiert hat, während in den beiden dazwischenliegenden Generationen sich keinerlei Anklänge dafür finden, ist *rezessive Vererbung wahrscheinlich.* Für das *Bettmann*sche Muster gilt übrigens dasselbe.

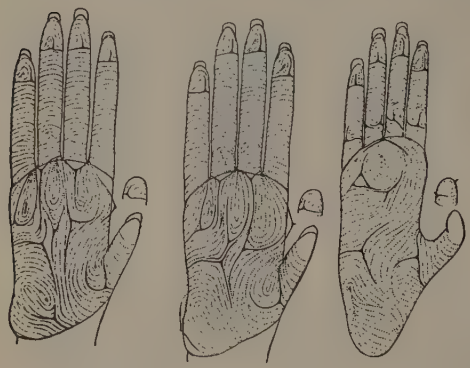
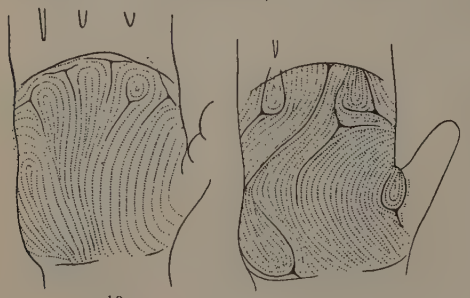


Abb. 17. Drei Hände von Pongo (Orang Utan). Alle Wirbel in J III und J IV streben zur Schleifenform, in J II ist sie schon erreicht bei der ersten und zweiten Hand; die dritte trägt nur einen einzigen großen Wirbel in J III.



18 19

Abb. 18. Eine Gorilla-Hand mit einem einzigen Wirbel in J II. Dieser Typus kommt bei Menschen vor, ist aber selten, außerdem münden die Schleifen dann distal. *Abb. 19.* Eine Schimpansen-Hand. Es sind nur noch Schleifen vorhanden. Viele Menschen haben diesen Typus. Hier münden die Schleifen distal.

Zusammenfassung.

Primitive Wirbel innerhalb der Interdigitalräume der Menschenhand sind äußerst selten. Es wird ein Beispiel einer Familie gegeben, wo bei Urgroßmutter und Urenkeltochter jene Wirbel, wenn auch zum Teil progressiv abgewandelt,

gefunden wurden. Man kann daher, phylogenetisch gesehen, vom «Makaktypus» sprechen.

Summary.

Primitive whorls in the interdigital spaces are extremely rare in the hands of man. An example is given where great grandmother and great grandchild bear such whorls. Therefore it is possible to speak of "Makak type" because that ape, seen phylogenetically, is the last one that bears so many whorls.

Résumé.

Des tourbillons primitifs au dedans des espaces interdigitaux de la main humaine sont extrêmement rares. On donne un exemple, où l'on a constaté ces tourbillons chez l'arrière-grand-mère et son arrière-petite-fille. C'est pourquoi on peut parler du «type Macaque», car ce singe, au point de vue phylogénétique, est le dernier qui porte tant de tourbillons.

LITERATUR

1. *Cummins, H., C. Keith, C. Midlo* et al.: Amer. J. phys. Anthrop. *12*, 415, 1929. – 2. *Midlo, C.* and *H. Cummins:* Palmar and Plantar Dermatolgyphics in Primates. Amer. anat. Mem. No. 20, Wistar Institute of Anatomy and Biology, Philadelphia 1942. – 3. *Wipple, I.:* Z. Morph. Anthrop. *7*, 261, 1904. – 4. *Wilder, H. H.* Biol. Bull. *30*, 1916; *49*, 1925.

Gini, C.: Acta genet. *4*, 175–191, 1953

THE MEASUREMENT OF THE DIFFERENCES BETWEEN TWO QUANTITY GROUPS AND IN PARTICULAR BETWEEN THE CHARACTERISTICS OF TWO POPULATIONS

By CORRADO GINI, Rome

In illustrating the difference between two quantity groups in general and between the measurements of a characteristic in two populations in particular, it is important to know not only what the difference is between the averages of the two quantity groups but also the degree of regularity with which the differences occur in individual cases in one direction rather than in another. For instance, in order to

realise the difference in stature between a group of males and a group of females, it is important not only to know the degree in which the average stature of the males exceeds the average stature of the females, but also the frequency with which the stature of a single male exceeds or falls below that of a single female. The difference between the averages gives us a synthetic idea of the *intensity* of the differences between the measurements of the characteristic in the two groups. The frequency with which, in individual cases, the difference occurs in the same direction or in the contrary direction to that of the averages, affords us a synthetic idea of the *typical character* of the difference between the averages.

When the difference between two individual cases is in an opposite direction to that between the averages of the two groups there is said to be *transvariation*.

A systematic treatment of this subject was made by the present writer as far back as 1916 [1]. Two transvariation indices were then proposed:

(1) – *the probability of transvariation*, given by the ratio of the number of the transvariations that occur between the two groups to the maximum number that might occur;

(2) – *the intensity of transvariation* given by the ratio of the sum of the transvariations to the maximum that the said sum may attain.

These two indices of transvariation were largely used in Italy, above all by *Boldrini* and his pupils, for measuring the regularity of differences between males and females in respect of multiple characteristics [2].

Subsequently, various other indices of transvariation were considered.

(3) – The *transvariability* that is given by the relation of the number of transvariations to the number of the differences that can be established between the quantities of the two groups [3].

(4) – *The area of transvariation* that is given by the area common to the frequency curves of the two characteristics equated to the sum of the areas of the two curves [4].

(5) – *The relative area of transvariation* which is equal to the area of transvariation equated to its maximum [4].

(6) – *The ratio of transvariation* which is equal to the relative area of transvariation when the two distributions contain an equal number of cases [4].

(7) – *The error of the discriminative value*, calling discriminative the value common to the two distributions which minimises the error committed by admitting that all the values of the group with the higher average are greater, and all those of the group with the lower average are inferior to the said discriminative value [4].

(8) – *The error of critical value*, calling critical the discriminative value in the special case in which the two distributions contain an equal number of cases [5].

(9) – *The area of security* which is given by the percentage of cases in which the value is found only in one or only in the other of the two distributions.

(10) – *The ratio of security* which is given by the area of security in the special case in which the two distributions contain an equal number of cases [6].

The first 8 indices are *direct indices of transvariation* and therefore inverse indices of the typical character of the differences between the averages of the two groups, while the last two are inverse indices of transvariation, and consequently direct indices of the typical character of the said differences.

It should be noted that the indices 7, 8, 9, 10, can be calculated both for the two groups of quantities as a whole, and separately for each of them.

The indices of transvariation mentioned above have been at first applied to individual characteristics of two populations (*monodimensional indices*), so that, should one want to compare the multiple characteristics of two populations, it would be necessary to calculate as many monodimensional indices as there were characteristics considered. Now it is important, in such a case, to be able to summarise the comparison in a synthetic index (a *pluri-dimensional index*) which takes into account the differences between the two populations with respect to all characteristics considered [7].

It was *Dahlberg* who in 1941 showed a method suited to this purpose which he has also applied to the differences existing between Swedes and Lapps as regards stature, the cephalic index and the facial index, determining the monodimensional, bidimensional, and tridimensional ratios of security for these characteristics [8].

Dahlberg's method is based on the supposition that the distributions of the individual characteristics considered are normal, and that the correlation between the average measurements of these characteristics in the two groups has the same sign as that of the

correlation between the individual measurements of each group, a supposition which holds good for the material relating to Swedes and Lapps which he elaborated, but which may not occur in other cases.

In the following pages another method for obtaining a synthetic index of the typical character of the differences between the characteristics of the two groups [9] is set forth, a method which is independent of any hypothesis concerning either the distribution of the individual characteristics or the relations that exist between them.

It will also be applied to the data already elaborated by *Dahlberg*, which he has kindly communicated to me, obtaining all the indices of monodimensional, bidimensional, and tridimensional transvariations of the ten types mentioned above for the three characteristics in question.

As will be seen, our conclusions agree except on one point of which we shall speak and which perhaps deserves further enquiries.

* * *

The premise for the determination of a pluridimensional index of transvariation is the ability to measure the importance of the difference between the measurements of a given characteristic in two groups in respect of the importance of the difference between the measurements of a second characteristic.

If, for instance, a woman presents a higher thoracic index but a lower stature than a man, should we say that there is or is not transvariation as regards these two characteristics taken as a whole? The reply will depend, of course, on the relative degree of intensity of the difference in relation to the two characteristics. If the woman is much shorter than the man and if, on the other hand, her thoracic index is only a little higher, we should be justified in saying that there is not transvariation; but if the woman is only a little shorter than the man and has a markedly higher thoracic index, the contrary affirmation would be justified. In order to solve the question, it is necessary to define what is the difference between the statures in one direction that can be taken as compensating a certain difference between the thoracic indices in the other direction.

Now it is clear that greater importance should be given to the difference between the measurements of a characteristic in a pair of cases belonging to two different groups, the greater the absolute value of the difference in question, and that less should be given the greater is the average $\triangle_{1,\,2}$ between all the differences which can be

established between the measurements of the characteristic in the two groups. Thus, if a woman of a given population differs in stature by one unit of measurement (say, for instance, a centimeter or an inch) from the stature of a man, it is evident that the difference will have only half the importance it would have if it were by two units of measurements; but, on the other hand, if a woman differs in stature from a man by one unit of measurement, the difference will be ten times less important if the differences which can be established between the statures of the woman and those of the man average one unit of measurement than if they average ten such units; for in the first case the difference between the two individuals considered is of the same size as the average of all the possible differences, whereas, in the second case, it is only one tenth of that average.

Generally speaking, if the two groups be termed 1 and 2, let m_1 stand for the number of individuals in group 1, m_2 for the number of the individuals in group 2, and therefore $m_1 \cdot m_2$ for the number of different couples which can be formed with an individual selected at random from group 1 and another selected at random from group 2. $a_{1/i}$, $a_{2/i}$ stand for the measurements of the characteristic α acquired in the two individuals of the i^{ma} couple ($i = 1, 2, \ldots m_1 \cdot m_2$) and similarly $b_{1/i}$, $b_{2/i}$; $c_{1/i}$, $c_{2/i}$; .... $n_{1/i}$, $n_{2/i}$ for the measurements of the characteristics β, γ, ... ν in the same individuals of the i^{ma} couple. Hence $a_{1/i}$—$a_{2/i}$, $b_{1/i}$—$b_{2/i}$, $c_{1/i}$—$c_{2/i}$, ... $n_{1/i}$—$n_{2/i}$ are the respective differences. ${}_{\alpha}\triangle_{1,2}$, ${}_{\beta}\triangle_{1,2}$, ${}_{\gamma}\triangle_{1,2}$ .... ${}_{\nu}\triangle_{1,2}$ stand for the averages of the $m_1 \cdot m_2$ differences, which can thus be established for the $m_1 m_2$ couples between the measurements of the α, β, γ, ... ν characteristics. The said differences $a_{1/i}$—$a_{2/i}$, $b_{1/i}$—$b_{2/i}$, $c_{1/i}$—$c_{2/i}$ .... $n_{1/i}$—$n_{2/i}$ will be comparable one with the other when divided by their respective mean differences ${}_{\alpha}\triangle_{1,2}$, ${}_{\beta}\triangle_{1,2}$, ${}_{\gamma}\triangle_{1,2}$ .... ${}_{\nu}\triangle_{1,2}$.

The ratios, comparable with one another, thus obtained

$$\frac{a_{1/i} - a_{2/i}}{{}_{\alpha}\triangle_{1,2}}, \quad \frac{b_{1/i} - b_{2/i}}{{}_{\beta}\triangle_{1,2}}, \quad \frac{c_{1/i} - c_{2/i}}{{}_{\gamma}\triangle_{1,2}}, \quad \ldots \quad \frac{n_{1/i} - n_{2/i}}{{}_{\nu}\triangle_{1,2}}$$

will be known as *reduced differences*.

It is easy to convince oneself that if, for the purpose of determining the probability or the intensity or any other monodimensional index of transvariation, we work on the reduced differences, we shall obtain the same results as would be obtained if we operated on the original differences.

* * *

If now we sum up for each couple the respective reduced differences relating to 2, 3 ..... n characteristics, we shall obtain quantities which we will describe as *composite bidimensional, tridimensional, n-dimensional composite differences.* With these we shall be able to work for the purpose of obtaining the probability or the intensity or other bidimensional, tridimensional .... n-dimensional indices of transvariation.

For the couple i^{ma}

$$ {}^{1}_{\alpha}D_{1,\,2/i} = \frac{a_{1/i} - a_{2/i}}{{}_{\alpha}\triangle_{1,\,2}}, \quad {}^{1}_{\beta}D_{1,\,2/i} = \frac{b_{1/i} - b_{2/i}}{{}_{\beta}\triangle_{1,\,2}}, \quad \ldots\ldots \quad {}^{1}_{\gamma}D_{1,\,2/i} = \frac{n_{1/i} - n_{2/i}}{{}_{\gamma}\triangle_{1,\,2}} $$

will indicate *the reduced differences;*

$$ {}^{2}_{\alpha,\,\beta}D_{1,\,2/i} = \frac{a_{1/i} - a_{2/i}}{{}_{\alpha}\triangle_{1,\,2}} + \frac{b_{1/i} - b_{2/i}}{{}_{\beta}\triangle_{1,\,2}}; \quad {}^{2}_{\alpha,\,\gamma}D_{1,\,2/i} = \frac{a_{1/i} - a_{2/i}}{{}_{\alpha}\triangle_{1,\,2}} + \frac{c_{1/i} - c_{2/i}}{{}_{\gamma}\triangle_{1,\,2}} \ldots\ldots $$

$$ \ldots\ldots {}^{2}_{\gamma,\,\nu}D_{1,\,2/i} = \frac{c_{1/i} - c_{2/i}}{{}_{\gamma}\triangle_{1,\,2}} + \frac{n_{1/i} - n_{2/i}}{{}_{\nu}\triangle_{1/i}} $$

the composite bidimensional differences;

$$ {}^{3}_{\alpha,\,\beta,\,\gamma}D_{1,\,2/i} = \frac{a_{1/i} - a_{2/i}}{{}_{\alpha}\triangle_{1,\,2}} + \frac{b_{1/i} - b_{2/i}}{{}_{\beta}\triangle_{1,\,2}} + \frac{c_{1/i} - c_{2/i}}{{}_{\gamma}\triangle_{1,\,2}}, \quad \ldots\ldots $$

$$ \ldots\ldots {}^{3}_{\alpha,\,\beta,\,\nu}D_{1,\,2/i} = \frac{a_{1/i} - a_{2/i}}{{}_{\alpha}\triangle_{1,\,2}} + \frac{b_{1/i} - b_{2/i}}{{}_{\beta}\triangle_{1,\,2}} + \frac{n_{1/i} - n_{2/i}}{{}_{\nu}\triangle_{1,\,2}}, \quad \ldots\ldots $$

$$ \ldots\ldots {}^{3}_{\beta,\,\gamma,\,\nu}D_{1,\,2/i} = \frac{b_{1/i} - b_{2/i}}{{}_{\beta}\triangle_{1,\,2}} + \frac{c_{1/i} - c_{2/i}}{{}_{\gamma}\triangle_{1,\,2}} + \frac{n_{1/i} - n_{2/i}}{{}_{\nu}\triangle_{1,\,2}} \ldots\ldots $$

the *composite tridimensional differences*, and so forth; finally

$$ {}^{n}_{\alpha,\,\beta,\,\ldots\,\nu}D_{1,\,2/i} = \frac{a_{1/i} - a_{2/i}}{{}_{\alpha}\triangle_{1,\,2}} + \frac{b_{1/i} - b_{2/i}}{{}_{\beta}\triangle_{1,\,2}} + \frac{c_{1/i} - c_{2/i}}{{}_{\gamma}\triangle_{1,\,2}} + \ldots\ldots + \frac{n_{1/i} - n_{2/i}}{{}_{\nu}\triangle_{1,\,2}} $$

the composite n-dimensional difference.

It should be noted that in adding up the reduced differences for the purpose of obtaining the composite differences, each reduced difference should be given a plus or a minus sign depending on whether it has the same or a different sign from that of the difference between the corresponding averages of the two characters in groups 1 and 2 [10].

* * *

But by operating in this way it would be necessary to determine the $m_1 \cdot m_2$ differences, and then to calculate $m_1 \cdot m_2$ divisions so as to obtain the corresponding reduced difference, and this would make the application of the method prohibitive.

The application becomes however feasible if we have recourse to the notion of *reduced value.*

For this purpose it should be noted that the formula of a composite difference (let us take, for instance, the *composite n-dimensional difference*) may be written as follows:

$$_{\alpha,\beta,\gamma\ldots\nu}{}^{n}D_{1,2/i} = \frac{a_{1/i}}{{}_{\alpha}\triangle_{1,2}} + \frac{b_{1/i}}{{}_{\beta}\triangle_{1,2}} + \frac{c_{1/i}}{{}_{\gamma}\triangle_{1,2}} + \ldots + \frac{n_{1/i}}{{}_{\nu}\triangle_{1,2}} -$$

$$- \frac{a_{2/i}}{{}_{\alpha}\triangle_{1,2}} - \frac{b_{2/i}}{{}_{\beta}\triangle_{1,2}} - \frac{c_{2/i}}{{}_{\gamma}\triangle_{1,2}} - \ldots - \frac{n_{2/i}}{{}_{\nu}\triangle_{1,2}}$$

in which each of the values $a_{1/i}$, $b_{1/i}$, $c_{1/i}\ldots n_{1/i}$ and respectively $a_{2/i}$, $b_{2/i}$, $c_{2/i}\ldots n_{2/i}$ must be taken with a plus or minus sign depending on whether the respective average of the characteristic in its own group is higher or lower than that of the other group.

The single values:

$$_{\alpha}V_1 = \frac{a_{1/i}}{{}_{\alpha}\triangle_{1,2}};\ {}_{\beta}V_1 = \frac{b_{1/i}}{{}_{\beta}\triangle_{1,2}};\ \ldots\ {}_{\nu}V_1 = \frac{n_{1/i}}{{}_{\nu}\triangle_{1,2}};$$

$$_{\alpha}V_2 = \frac{a_{2/i}}{{}_{\alpha}\triangle_{1,2}},\ \ldots\ {}_{\nu}V_2 = \frac{n_{2/i}}{{}_{\nu}\triangle_{1,2}}$$

will be called *reduced values* and the values deduced from the sums of two, three .... *n* reduced values, relating to the same individual will be called *composite bidimensional, tridimensional .... n-dimensional values;* thus, for instance, the composite bidimensional values relating to the individual i^{mo} of group 1 will be

$$_{\alpha,\beta}{}^{2}V_{1/i} = \frac{a_{1/i}}{{}_{\alpha}\triangle_{1,1}} + \frac{b_{1/i}}{{}_{\beta}\triangle_{1,2}};\ \ldots\ {}_{\beta,\gamma}{}^{2}V_{1/i} = \frac{b_{1/i}}{{}_{\beta}\triangle_{1,2}} + \frac{n_{1/i}}{{}_{\gamma}\triangle_{1,2}};\ \ldots$$

the composite tridimensional values, always referring to the same individual, will be

$$_{\alpha,\beta,\gamma}{}^{3}V_{1/i} = \frac{a_{1/i}}{{}_{\alpha}\triangle_{1,2}} + \frac{b_{1/i}}{{}_{\beta}\triangle_{1,2}} + \frac{c_{1/i}}{{}_{\gamma}\triangle_{1,2}};\ \ldots$$

$$_{\alpha,\beta,\nu}{}^{3}V_{1/i} = \frac{a_{1/i}}{{}_{\alpha}\triangle_{1,2}} = + \frac{b_{1/i}}{{}_{\beta}\triangle_{1,2}} + \frac{n_{1/i}}{{}_{\gamma}\triangle_{1,2}};\ \ldots$$

and the composite n-dimensional value for this individual will be

$$_{\alpha,\beta,\gamma\ldots\nu}{}^{n}V_{1/i} = \frac{a_{1/i}}{{}_{\alpha}\triangle_{1,1}} + \frac{b_{1/i}}{{}_{\beta}\triangle_{1,2}} + \frac{c_{1/i}}{{}_{\gamma}\triangle_{1.2}} + \ldots + \frac{n_{1/i}}{{}_{\nu}\triangle_{1.2}}$$

In this way, the calculations are greatly abbreviated because, instead of calculating $m_1 \cdot m_2$ reduced differences, it will be enough to calculate $m_1 + m_2$ reduced values.

* * *

The preliminary calculations of a pluri-dimensional index of transvariation are therefore reduced to the following:

1. Calculate the n *average differences*

$$_{\alpha}\triangle_{1,2},\ _{\beta}\triangle_{1,2},\ _{\gamma}\triangle_{1,2},\ \cdots\ _{\nu}\triangle_{1,2}$$

A formula for the calculation of these average differences has been given by the author of this paper and can be found on page 153 of his *Memorie di metodologia statistica*, Giuffrè, Milan 1935 [11].

2. Calculate for each of the n characteristics the m_i *reduced values* relating to the m_1 individuals of the group 1 and the m_2 *reduced values* relating to the m_2 individuals of group 2.

3. For each of the m_i individuals of group 1 and each of the m_2 individuals of group 2, add up the reduced values, thus obtaining the respective *pluri-dimensional composite values.*

We already know how, by operating on the m_1 values of group 1 and on the m_2 values of group 2 relating to a characteristic, we obtain the probability, the intensity, and the other *monodimensional indices of transvariation.* By operating similarly on the m_1 composite pluri-dimensional values of group 1 and on the m_2 composite pluri-dimensional values of group 2 relating to more characteristics, we obtain the respective *pluri-dimensional indices of transvariation* for the said characteristics.

* * *

Let us see how the above-stated method can be applied in practice.

We will examine, for this purpose, the data relating to the stature, the cephalic index and the facial index of 1000 Swedes and 269 Lapps which have been worked out in the article by Prof. *G. Dahlberg* referred to above, and which he has kindly communicated to me.

For the sake of brevity, we reproduce in Table 1 only the data relating to the first three and the last three Swedes, and those relating to the first 3 and the last 3 Lapps.

The data for the Swedes have been taken from the cards of the Hollerith machines which give only full numbers. They should therefore be taken as abbreviated: for instance, three statures of cm 179.0, 179.4, 179.9 are all entered as equal to cm 179.

For the sake of uniformity, the same rule has been followed for the data relating to the Lapps.

The arithmetical averages for the three characteristics thus measured are found to be:

	Stature (cm)	Cephalic index (%)	Facial index (%)
Swedes	171.8	77.1	92.5
Lapps	158.5	82.8	82.5

which, for the circumstance noted above, should be augmented by 0.5 to obtain the actual average values.

Table I. Original Values.

Rank	Stature (cm) (a_1)	Cephalic index (%) (b_1)	Facial index (%) (c_1)
	Swedes		
1	167	79	91
2	174	78	83
3	167	74	90
.....			
.....			
998	167	79	91
999	167	79	91
1000	173	78	98
	Lapps		
	(a_2)	(b_2)	(c_2)
1	138.5	88.57	82.68
2	140.0	84.86	85.52
3	143.0	87.22	75.00
.....			
.....			
267	172.5	84.95	82.88
268	173.5	84.46	75.32
269	179.2	81.96	79.08

The average differences between the original measurements of the Swedes and of the Lapps in all their possible combinations are found to be, for the three characteristics considered,

Stature (cm)	Cephalic index (%)	Facial index (%)
$_{\alpha}\triangle_{1,2} = 13.71$	$_{\beta}\triangle_{1,2} = 6.25$	$_{\gamma}\triangle_{1,2} = 10.82$

If we divide the respective original measurements given in Table 1 by those of the aforesaid average differences, we obtain the *reduced values* shown in Table II.

From these the *pluri-dimensional composite values* are obtained. To obtain them, the reduced values of the statures and of the facial indices—for which the averages are higher in the case of the Swedes—must be taken with a plus sign in the case of the latter and with a minus sign in the case of the Lapps, while the reduced values for the cephalic indices (the average for which is higher in the case of the Lapps) must be taken with a plus sign for the latter and with a minus sign for the Swedes.

Table II. Reduced values.

Rank	Stature (cm) (αV_1)	Cephalic index (%) (βV_1)	Facial index (%) (γV_1)
	Swedes		
1	12.18	12.64	8.41
2	12.69	12.48	7.67
3	11.67	11.84	8.32
.....	.....	.....	.....
.....	.....	.....	.....
998	12.18	12.64	8.41
999	12.18	12.64	8.41
1000	12.62	12.41	9.06
	Lapps		
	(αV_2)	(βV_2)	(γV_2)
1	10.06	14.08	7.58
2	10.21	13.44	7.86
3	10.43	13.92	6.93
.....	.....	.....	.....
.....	.....	.....	.....
257	12.54	13.44	7.58
258	12.62	13.44	6.93
259	13.05	12.96	7.30

The reduced values of Table II (or, which amounts to the same thing, the original measurements of Table I), and the composite values of Table III, have been used for calculating, by the methods known, the several indices of transvariation, as well as the discriminative value and the critical value, thus obtaining the results set forth in Table IV.

Table IV omits the values of the transvariability which, in our case, are found to be always equal to half of the values of the probability of transvariation.

The values of the errors of the discriminative value and of the critical value have an importance of their own only in so far as they refer separately to the Lapps (columns 6 and 11) and to the Swedes (columns 7 and 12), because in so far as the total values are concerned (columns 5 and 10) they are in this case equivalent (apart from the approximations of the calculations affecting the last decimal figures) to half of the values of the relative area of transvariation (column 3) and respectively to the ratio of transvariation (column 8).

It should be borne in mind that while all the other indices (column 3) and respectively to the ratio of transvariation (column 8). area and the ratio of security (columns 13–16) are inverse indices of transvariations.

Table III. Pluri-dimensional composite values

Rank	Bidimensional composite values			Tridimensional composite values
	Stat.-ceph.i.	Stat.-fac.i.	Ceph.i.-fac.i.	Stat.-ceph.i.-fac.i.
	$(\alpha V_1 - \beta V_1)$	$(\alpha V_1 - \gamma V_1)$	$(\gamma V_1 - \beta V_1)$	$(\alpha V_1 + \gamma V_1 - \beta V_1)$
		Swedes		
1	–0.46	20.59	–4.23	7.95
2	0.21	20.36	–4.81	8.88
3	–0.17	19.99	–3.52	8.15
.....				
.....				
998	–0.46	20.59	–4.23	7.95
999	–0.46	20.59	–4.23	7.95
1000	0.21	21.68	–3.35	9.27
		Lapps		
	$(\beta V_2 - \alpha V_2)$	$(\alpha V_2 - \gamma V_2)$	$(\beta V_2 - \gamma V_2)$	$(\beta V_2 - \alpha V_2 - \gamma V_2)$
1	4.02	–17.64	6.50	–3.56
2	3.23	–18.06	5.58	–4.63
3	3.49	17.36	6.99	–3.44
.....				
.....				
267	0.90	–20.12	5.86	–6.68
268	0.82	–19.55	6.51	–6.11
269	–0.99	–20.35	5.66	–7.39

The examination of the mono-dimensional indices of transvariation shows that the transvariation is lowest—and therefore the typical character of the difference between the averages is highest—for stature, followed by the facial index and lastly by the cephalic index.

This conclusion is supported by all the indices: probability (column 1), intensity (column 2), area (column 3), ratio (column 8) of transvariation, total error of the discriminative value (column 5), total error (column 9) of the critical value, area (column 13) and ratios (columns 14, 15, 16) of security. On the other hand, it is not shown—and it was not to be expected that it would show—for the errors of the discriminative value nor for the errors of the critical value calculated separately for the Lapps (columns 6 and 11) and for the Swedes (columns 7 and 12).

As a result of the fact that the number of Lapps (269) is smaller than that of the Swedes (1000), the values of the critical error are lower for the Lapps (columns 11 and 6) and higher for the Swedes (see columns 12 and 7) of the corresponding errors of the discriminative value.

Table IV. Indices of transvariation of the distribution of 1000 Swedes and 269 Lapps according to three anthropometric characteristics (Stature, cephalic index, facial index).

Characteristics	P	I	A	V.D.	E.D.	E.D.L.	E.D.S.	r	V.C.	E.C.	E.C.L.	E.C.S.	A S	R S	R S L	R S S
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Monodimensional indices																
Stature . . .	0.120	0.0605	0.199	162.10	0.0993	0.249	0.059	0.270	164.62	0.135	0.164	0.106	0.366	0.347	0.312	0.381
Cephalic index .	0.233	0.149	0.260	81.64	0.136	0.361	0.0751	0.359	79.79	0.179	0.160	0.199	0.048	0.051	0.055	0.046
Facial index .	0.201	0.135	0.254	85.11	0.127	0.245	0.095	0.322	86.58	0.161	0.186	0.136	0.078	0.082	0.089	0.075
Bidimensional indices																
Stat.-ceph. ind.	0.148	0.0928	0.194	1.29	0.0969	0.290	0.045	0.276	0.92	0.139	0.152	0.125	0.093	0.138	0.215	0.061
Stat.-fac. ind. .	0.0634	0.0431	0.126	19.68	0.0630	0.0234	0.017	0.185	20.13	0.0929	0.097	0.089	0.494	0.491	0.483	0.498
Ceph. ind.-fac. ind.	0.118	0.0713	0.180	4.99	0.0906	0.238	0.051	0.225	4.67	0.114	0.108	0.121	0.294	0.240	0.144	0.335
Tridimensional indices																
Stat. ceph. ind. fac. Ind. . .	0.0532	0.0275	0.101	6.87	0.0504	0.171	0.018	0.152	7.52	0.0820	0.0629	0.101	0.390	0.427	0.490	0.363

P = Probability of transvariation
I = Intensity of transvariation
A = Relative area of transvariation
r = Ratio of transvariation
VD = Discriminative value
VC = Critical value. (The figures of monodimensional indices for VD and VC are original values, those of pluridimensional indices, reduced values.)
ED = Error of the total DV
EDL = Error of the DV for the Lapps
EDS = Error of the DV for the Swedes
EC = Error of the total CV
ECL = Error of the CV for the Lapps
ECS = Error of the CV for the Swedes
AS = Area of security
RS = Ratio of the security
RSL = Ratio of security for the Lapps
RSS = Ratio of security for the Swedes. (For the same group—Lapps or Swedes—the area of security and the ratio of security coincide.)

If we compare the mono-dimensional indices with the bi-dimensional ones, and these with the tri-dimensional, we see that often, though not always, the bi-dimensional indices show a lower transvariation of both the mono-dimensional indices concerning the two characteristics to which the bi-dimensional index refers, and almost always—though not quite always—the tri-dimensional indices show a lower transvariation of the three bi-dimensional indices concerning the three couples of the characteristics which are considered in the tri-dimensional index.

Among the bi-dimensional indices, the probability, the intensity and the ratio of transvariation, the total error of critical value, the error of critical value for the Swedes, the total area and the total ratio of security, all show, for the stature-cephalic index couple, a higher degree of transvariation than is obtained by the respective mono-dimensional indices for stature.

In the case of the tri-dimensional index, the total area and the total ratio of security are found to be lower—i.e. the transvariation is found to be higher—than the respective bi-dimensional indices for stature and the facial index.

Therefore it is found that when a larger number of characteristics are considered, transvariation does not always diminish and the typical character of the difference between the averages of the two groups does not consequently always increase.

* * *

Instead of determining, as has been done in the preceding pages, the actual values of the indices of transvariation, we might have ascertained their respective theoretical values on the assumption of the normal distribution of the original and composite values, by applying the formulae given in the paper: *Nuovi contributi alla teoria della transvariazione* (*New Contributions to the Theory of Transvariation*) presented by Dr. *Livada* and the present writer to the meeting of the Italian Society of Statistics, on June 28, 1943 [12].

Whether it is advisable to determine only the real values or also the theoretical values of the indices of transvariation will depend on whether the two groups of quantities are more or less numerous—and their distribution consequently more or less subject to the influences of chance—and on whether the aforesaid distributions vary more or less from the normal.

* * *

In his paper of 1941, *Dahlberg* calculates mono-dimensional and pluri-dimensional indices of transvariation belonging to those which we have called ratios of security, on the hypothesis that the distributions of the characteristics follow the normal curve, and he comes to the conclusion that both for the Swedes and for the Lapps, the aforesaid mono-dimensional ratios for stature are lower than the bi-dimensional ratios for stature and cephalic index, and that the bi-dimensional indices for stature and facial index are lower than the tri-dimensional ratio for stature, cephalic index, and facial index. This differs from the results obtained by us and set forth in columns 14, 15 and 16 of Table IV.

This discordance may in part depend on the fact that *Dahlberg* has based his calculations on the normal rather than on the actual distributions.

It may also certainly depend in part on the fact that *Dahlberg's* method for obtaining the pluri-dimensional indices differs from ours.

But, apart from these circumstances, there is a more intrinsic reason which should be studied thoroughly.

Dahlberg, indeed, affirms that the larger the number of characteristics considered, the larger the ratio of security, that is to say that transvariation diminishes, when the number of characteristics is sufficient, until transvariation disappears. Now, it does not seem that on general lines this affirmation can be accepted. It would, for instance, seem natural that if, in the case of two groups and with regard to a certain characteristic, transvariation is absent or is minimum, and, besides the characteristic in question, another be considered in which transvariation is marked, the bi-dimensional indices of transvariation should be higher and cannot be lower than the mono-dimensional indices of the first characteristic in question, as we have indeed ascertained for most of the bi-dimensional indices of stature and cephalic index as compared to the respective mono-dimensional indices of stature.

This is a point of the theory which deserves to be made the subject of a thorough study.

[1] See *Il concetto di «transvariazione» e le sue prime applicazioni*, in «Giornale degli Economisti», January 1916, reproduced in *Memorie di Metodologia Statistica*, Milan 1939, pp. 473–527 with several additions and an Appendix *Sul massimo della probabilità di transvariazione e della intensità di transvariazione tra due gruppi.*

[2] *M. Boldrini, Su alcune differenze sessuali secondarie nelle dimensioni del corpo umano alla nascita e alle età superiori*, «Archivio per l'Antropologia e la Etno-

logia», Vol. XLIX, 1919; *I cadaveri degli sconosciuti, Ricerche demografiche ed antropologiche della Morgue di Roma*, «La Scuola Positiva», 1920; *Gli studi statistici sul sesso: la proporzione dei sessi nelle nascite e i caratteri sessuali secondari.* «Rassegna di studi sessuali», I, 1921; *Differenze sessuali nei pesi del corpo e degli organi umani*, «Rendiconti della R. Accademia dei Lincei, Classe di Sc. Fis. Mat. e Nat.», Vol. XXIX, series 5°, 2° sem. Fasc. 1–4, Rome 1920; *Misure interne ed esterne di alcune ossa lunghe nell'uomo e nella donna*, Ibidem, vol. XXXII, series 5°, 2° sem. Fasc. 1–10, Rome 1924, pp. 233–253; *F. Giglio, Differenze sessuali nei caratteri quantitativi degli embrioni umani*, «Atti del Congresso internazionale per gli studi sulla popolazione», Rome 1931, Vol. II, pp. 533–543.

See also the article by *D. Miani Calabrese, La transvariazione rispetto al sesso dei caratteri fisici dell'infanzia*, in «Statistica», October-December 1943.

A chapter on transvariation is contained in our *Corso di statistica* (see the second Spanish edition, published by the Editorial Labor S. A. Barcelone, 1953 and the more recent Italian edition, edited by Dr. *S. Gatti* and Dr. *C. Benedetti*, University of Rome, Ac. year 1952–53, Rome, Veschi), as well as in the treatises of *M. Boldrini, Biometrica. Problemi della vita delle specie e degli individui*, Cedam Padua 1927; *Statistica, Teoria e metodi*, Milan, Giuffré, Ist edition, 1942, IInd edition, 1950; *L. Galvani, Lezioni di statistica metodologica*, Guf, Neapel 1935; *G. Barbensi, Elementi di metodologia biometrica*, Niccolai, Florence 1940.

On the standard deviation of the probability of transvariation, see the articles by *V. Castellano, Sullo scostamento quadratico medio della probabilità di transvariazione*, «Metron», Vol. XI, No. 4, 1934, pp. 19–75 and *G. Ottaviani, Sulla probabilità che una prova su due variabili casuali X e Y verifichi la disuguaglianza $X<Y$ e sul corrispondente scarto quadratico medio*, «Giornale dell'Istituto Italiano degli Attuari», Year X, No. 3–4, July-October 1939, pp. 185–192.

A suitable method for the calculation of the intensity of transvariation has been given by Dr. *G. Livada* in the paper *Procedimento per il calcolo della intensità di transvariazione*, in «Atti della VI e VII Riunione scientifica, Roma, Gennaio 1943, Giugno 1943», Italian Society of Statistics, Rome, pp. 63–71.

[3] This index is employed in the papers presented to the Italian Society of Statistics by Dr. *Giacomo Sonnino, Contributi alla teoria della transvariazione tra seriazioni correlate*, XI Riunione scientifica, April 1951 and by Dr. *Luigi de Lucia, Transvariazione tra caratteri connessi*, XII Riunione scientifica, February 1952. The volume of the Proceedings is under press.

[4] See the paper by *C. Gini* and *G. Livada, Nuovi contributi alla teoria della transvariazione*, in «Atti della VI–VII Riunione scientifica, Roma gennaio e giugno 1943», Rome, Italian Society of Statistics, 1945, pp. 346–371. For the ratio of transvariation, see also the monograph *Metodologia statistica. La misura dei fenomeni collettivi*, in «Enciclopedia delle matematiche elementari», Vol. III, Part. II, Hoepli, Milan 1948, pp. 296–299.

[5] See the paper by *Gini* and *Livada* cited in note [4] and that by *F. de Helguero, Il valore delle differenze sessuali dal punto di vista biometrico*, in «Atti della Società Romana d'Antropologia», Vol. XIII, Fasc. I, 1907, pp. 88–96.

[6] It belongs to ratios of security, the index proposed and applied by Professor *Dahlberg* in the paper cited in the following note [8].

[7] See our paper *Della misura sintetica della transvariazione rispetto ad n caratteri* (*transvariazione n-dimensionale*) presented to the Italian Society of Sta-

tistics at the XI Scientific meeting of April 1951. The volume of the Proceedings is under press. See also the paper (in collaboration with Dr. *G. Livada*) entitled *Transvariazione a più dimensioni* in «Atti della VI e VII Riunione Scientifica della Società Italiana di Statistica», January 1943, pp. 25–62.

[8] In the paper entitled *The conception of race and a method of delimiting races demonstrated on a material of Swedes and Lapps*, contained in the volume *The Race Biology of the Swedish Lapps*, Part II Uppsala 1941, pp. 32–41. The method has been summarised, with some particularizations, in the paper by us and *Livada* cited in the note [4]. Recently, Professor *Dahlberg* has made a more detailed presentation of the method in the German language in the article *Eine Methode zur Messung von Rassenunterschieden* published in «Metron», Vol. XVI, No. 1/2 (15 July 1951), pp. 133–150.

[9] Another generalization of the indices of transvariation leads to the measurement of the transvariation among several groups. See the paper *Per la determinazione della probabilità di transvariazione tra più gruppi*, in «Atti della IV e V Riunione Scientifica, Roma, Gennaio e Giugno 1942», Italian Society of Statistics, Rome 1942, pp. 195–197.

[10] When the two individual values of the characteristic in the two groups are equal, or when the averages of the characteristic in the two groups are equal, one cannot say whether there is or is not transvariation in regard to the said characteristic; in such a case it will be advisable to consider the differences for one half as positive and for the other half as negative. (See on the matter *Memorie di Metodologia statistica*, op. cit. p. 475.)

[11] Or else in *Indici di omofilia e di rassomiglianza* «Atti del R. Istituto Veneto di Scienze, Lettere ed Arti» Academic Year 1914–1915, Tomo IV, Part II. pp. 535–536, and in *Il concetto di «transvariazione» e le sue prime applicazioni*, «Giornale degli Economisti e Rivista di Statistica», Vol. LI, No. 1, January 1916.

[12] See «Atti della VI-VII Riunione Scientifica, Roma, Gennaio 1943 e Giugno 1943», Rome, Società Italiana di Statistica, 1945, pp. 346–371.

Summary.

The author calls attention to the concept of transvariation (which he subjected to a systematic analysis as early as in 1916) as well as to different indices which should make possible an adequate expression of this concept. These indices have generally been used to measure the so-called "typicité" of differences in regard to a certain characteristic between two population groups, for instance, to measure the regularity with which the stature is bigger in men as compared to women. It is, however, sometimes useful to have a common expression for the transvariation of several characteristics.

In 1941 *Dahlberg* proposed a method serving this purpose which he applied to the transvariation of height, cephalic index and facial index of Swedes and Lapps. *Dahlberg*'s method, however, was elaborated on the basis of certain hypotheses. In this report the author proposes another method which is independent of these hypotheses and has been applied to the same data. The different methods gave somewhat different results which should be subjected to a closer analysis.

Résumé.

L'auteur rappelle le concept de transvariation, dont il a donné un traité systématique dès 1916, ainsi que les différents indices proposés par lui-même et par d'autres auteurs pour en obtenir la mesure.

Ces indices sont en général appliqués pour mesurer la *typicité* des différences concernant un certain caractère dans deux groupes de population: par exemple, mesurer la régularité avec laquelle la taille est plus élevée parmi la population masculine que parmi la population féminine. Parfois il est pourtant intéressant de donner une mesure synthétique de la transvariation pour plusieurs caractères. *Dahlberg* a proposé une méthode qui répond à ce but dès 1941, et l'a appliquée à la mesure de la transvariation pour la taille, l'indice céphalique et l'indice facial entre Suédois et Lapons.

Cette méthode est pourtant subordonnée à certaines hypothèses. L'auteur propose ici une autre méthode qui en est indépendante et l'applique aux mêmes données.

Il attire l'attention sur une divergence entre les résultats des deux méthodes qu'il serait bon d'approfondir.

Zusammenfassung.

Der Verfasser erinnert an den Transvariationsbegriff, von dem er schon 1916 eine systematische wissenschaftliche Darstellung gegeben hat, sowie an verschiedene Indices, die von ihm und anderen Verfassern vorgeschlagen worden sind, um ein Maß für diesen Begriff zu erhalten.

Diese Indices werden im allgemeinen angewandt um «la typicité» in Unterschieden mit Rücksicht auf eine gewisse Eigenschaft bei zwei Bevölkerungsgruppen zu messen; um z. B. die Regelmäßigkeit zu messen mit der die Körperlänge unter der männlichen Bevölkerung größer ist als unter der weiblichen. Bisweilen ist es aber interessant, ein synthetisches Maß für die Transvariation mehrerer Eigenschaften zu geben. *Dahlberg* hat schon 1941 eine Methode vorgeschlagen, die diesen Zweck erfüllt, und hat diese auf Transvariationen der Körperlänge, Schädelindex und Gesichtsindex bei Schweden und Lappen angewandt.

Diese Methode beruht aber auf gewissen Hypothesen. Der Verfasser schlägt in vorliegender Arbeit eine andere Methode vor, die nicht auf solchen beruht und wendet diese auf die gleichen Aufgaben an.

Er verweist auf eine Abweichung in den Resultaten, die man bei Anwendung dieser verschiedenen Methoden erhält, und die man näher untersuchen sollte.

Glass, B. and *J. C. Kistler:* Acta genet. *4*, 192–206, 1953

DISTAL HYPEREXTENSIBILITY OF THE THUMBS

By BENTLEY GLASS and JOHN C. KISTLER

Department of Biology, The Johns Hopkins University

Among anatomical oddities, none would seem to be much more amusing or less consequential than doublejointedness at the terminal joint of the thumb (Fig. 1). This trait, technically termed distal hyperextensibility of the thumb, might more colloquially be called "hitchhiker's thumb." Although it has been previously mentioned in genetic literature as probably hereditary, no thorough investigation of the trait appears to have been made. *L. F. Whitney* [1932], in discussing doublejointedness at the proximal joint of the thumb, stated that he had also collected some data on the distal type of doublejointedness, and added: "This characteristic also runs in families, and it is interesting that it is inherited independently of the doublejointedness in the second joint." The projected paper, however, has never appeared. *D. D. Whitney*, in his book entitled Family Treasures, figures the condition in a mother and son (p. 157), and remarks that it is rather common and probably inherited (*D. D. Whitney* [1942], p. 163).

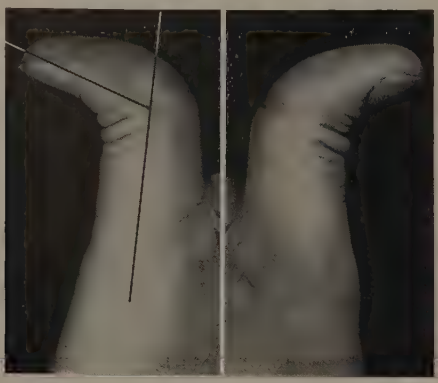


Fig. 1 A Fig. 1 B

Distal hyperextensibility of the thumbs does not appear to be affected by age or sex. The senior author, whose own thumbs are hyperextensible at the distal joints (see Fig. 1), can remember wondering about the peculiarity from his earliest childhood. In his father, at 77 years of age, it is equally marked. In spite of the commonness of the trait, however, most possessors of such thumbs suppose them rare, and are even surprised to find that other members of their own families are similarly characterized. Except in rare instances, persons with distally hyperextensible thumbs show no other type of doublejointedness. They are not in general loose-jointed.

X-ray photographs of normal thumbs were compared with those of distally hyperextensible thumbs from two individuals, one of whom had both thumbs and the other only a single thumb doublejointed in this characteristic way. It appeared that the distal end of the proximal phalanx is rather flatter and bears a sharper angle on the dorsal side, whereas the corresponding face of the proximal phalanx in a distally hyperextensible thumb exhibits a more rounded angle. More cases will have to be studied to substantiate such an interpretation, but it seems that the trait may have a skeletal basis and may not be due simply to longer ligaments.

The senior author's study of the trait in his own family (Fig. 2)

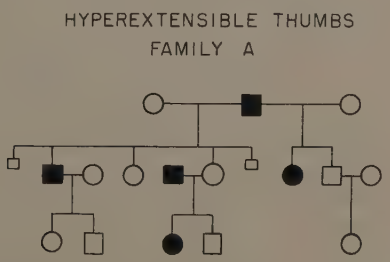


Fig. 2

indicated that distal hyperextensibility of the thumbs is hereditary. Its unbroken transmission might be taken for that of a simple autosomal dominant. However, its reintroduction into one branch of the family by marriage might imply that it is quite common. Examined from that point of view, it will be seen that there is no evidence from this pedigree to rule out the possibility that the trait is a simple autosomal recessive. In that case, a large proportion of the population

should be heterozygous, and the trait itself extremely common. A survey of several samples of the population was therefore undertaken.

The method used was the following. A transparent plastic protractor was reddened along the straight edge and two red lines were drawn parallel to this edge and respectively 0.5 and 1.0 cm from it. These assisted in aligning the straight edge of the protractor parallel to the posterior (dorsal) surface of the thumb, the hand being held palm upwards, and the midpoint of the protractor being placed on the end of the palmar skin folds, at the center of the distal joint of the thumb. The angle was then read along the line from this point to the corner or most laterally projecting part of the thumbnail (Fig. 1). All persons who had either thumb extensible to a 50° angle or greater were classified as possessors of the trait, all others as non-affected.

The first sample consisted of students at the Johns Hopkins University, who represented quite an extensive geographic area but came preponderantly from the Middle Atlantic States, with a large local contingent. This sample was compared with a more strictly local sample comprising mainly male students at a large high school in Baltimore, but including some adults of both sexes. Sibs and other close relatives were excluded. The third sample consisted of both male and female students at Morgan College in Baltimore. These represented the American Negro population. It soon became apparent that the trait is far commoner than had been supposed, and also, that some persons have only a single distally hyperextensible thumb, which may be either the right or the left. All persons with either thumb, or both thumbs, distally hyperextensible were considered in the enumeration to possess the trait. The frequencies found in the three samples of the population, and the penetrance, calculated on the assumption that non-expression of the condition in the right and left hands is independent, are given in Table 1. Approximately one-fourth of the white population and one-third of the Negro population examined consist of persons with distally hyperextensible thumbs. The penetrance is high. Only 3.5 per cent of those who carry the genotype for distally hyperextensible thumbs fail to manifest the trait on at least one hand or the other. The difference between the samples of the white population is not significant, but the frequency of the trait is significantly higher in the Negroes than in the Whites.

Following the original preparation of this paper and its presentation at the meeting of the Human Genetics Society in Minneapolis,

in September 1951, the paper of *Harris* and *Joseph* [1949] on the incidence of the same trait in 133 male and 100 female Europeans, 31 male Indians, and 30 male West Africans came to our attention. *Harris* and *Joseph* carried out their study by means of measurements on radiographs, and used the angle between the posterior surfaces of the proximal and distal phalanges as a standard measurement. They said that "sometimes the proximal phalanx does not show a well-defined straight line posteriorly, but in such cases a line may be drawn parallel to the greatest thickness of the cortical bone. The axis for the curved distal phalanx was more difficult to determine, since the phalanx is always curved in a posterior direction. However, there is nearly always a straight line of cortical bone about one millimetre wide on the posterior aspect, and a line drawn on the posterior surface of the phalanx parallel to this was chosen as the axis. It might be argued that the greater the curvature of the distal phalanx the greater should be the angle between it and the middle phalanx. It was found, however, that the curvature did not vary enough to affect the results materially." The angle was then measured to the nearest half degree. For the male Europeans the mean angle for the interphalangeal angle of the right thumb was 29.16° ($\sigma = 11.65^{\circ}$), and for the left thumb it was 32.20° ($\sigma = 11.09^{\circ}$). Graphs of the separate distributions for the two thumbs show fairly symmetrical continuous distributions, with a slight indication of a skewed or bimodal type of distribution.

This puzzling fact, in the light of the indication that the inheritance of the trait might be simple, led us to make a study (a) of the distribution of the interphalangeal angles from external measurements, and (b) of the comparative agreement between such measurements made independently by three observers on the same group of subjects. It was our hope that sufficient agreement and precision might be found in taking external measurements to obviate the need for measurements based on roentgenograms. Clearly, if the latter were essential, the trait would not be very useful to physical anthropologists and geneticists in field studies of human populations.

Measurements were made upon 122 male individuals, not the same persons as those in the earlier study. The method was the same as already described. The distributions obtained for right and left thumbs separately are shown in Figs. 3 and 4. The mean for the left thumb was $41.5^{\circ} \pm 0.85^{\circ}$ ($\sigma = 13.9^{\circ}$); and for the right thumb,

Fig. 3

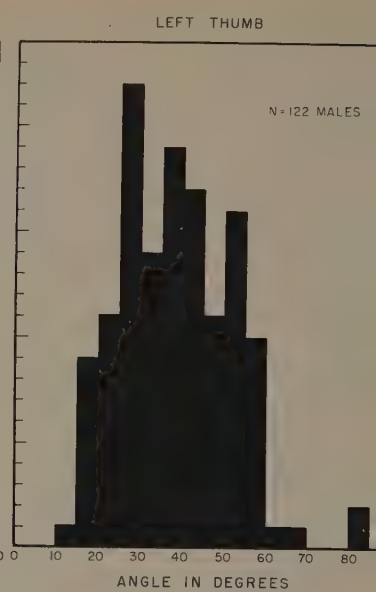


Fig. 4

$43.5^{\circ} \pm 0.825^{\circ}$ ($\sigma = 13.55^{\circ}$). It is interesting that although the median and range for the left thumb are higher than the median and range for the right, as in the data of *Harris* and *Joseph*, nevertheless the mean for the left thumb is lower than that for the right, contrary to the finding of *Harris* and *Joseph*. This is due to the asymmetry of the distributions. The differences between the means in the present series is not significant; and it is therefore to be presumed that the difference found by *Harris* and *Joseph* has no real meaning, either.

The difference between the means of the British workers' series and ours is 11.82°, owing to the different nature of the measurement. The ranges show the same general difference:

H. and *J.*	left	$5^{\circ} - 70^{\circ}$	right $0^{\circ} - 65^{\circ}$
G. and *K.*		$10^{\circ} - 85^{\circ}$	$0^{\circ} - 75^{\circ}$

The 50° angle taken as the criterion of distal hyperextensibility of the thumb would thus become about a 38° angle according to the standard measurement used by *Harris* and *Joseph*.

In our earlier study we classified each individual who had one

or both thumbs distally hyperextensible as a possessor of the trait, in contrast to persons who had neither thumb distally hyperextensible. It is unfortunately impossible from the manner in which *Harris* and *Joseph* have presented their data to make a similar classification of their subjects; although they did report a high correlation between right and left joints ($r = 0.66$ in males and 0.81 in females). When our new series of 122 male subjects was classified in the same way as the first, 30 (24.6 per cent) were found to have one thumb exceeding 50° in extension. This is in excellent agreement with the result obtained in the first series of 221 male subjects (see Table 1).

Table 1. A. Frequency of Persons with Hyperextensible Thumbs (dht) in Three Samples of the United States Population.

	dht	non-dht	Total	Percentage dht±S.E.
White: total	221	674	895	24.7±1.44
J.H.U.	85	294	379	22.4
B.C.	136	380	516	26.3
Negro: M.C.	56	101	157	35.6±3.8

χ^2 between two white samples, 1.82. P, .20 — .10
χ^2 between total white and negro, 8.36. P < .01

B. Penetrance of Hyperextensible Thumb (J.H.U. sample)

	Both	R. only	L. only	Total
	51	23	11	85
Frequency	.60	.27	.13	
Coincidence of non-expression		.27×.13 = .035		
Penetrance		100.0—3.5 = 96.5 %		

The three observers (*Ray A. Wilson*, *Lowell R. King*, and *Louis A. Mucelli*), to whom the authors acknowledge their indebtedness for the measurements on the second series of 122 male individuals, supplied a total of 726 deviations of measurements (3 observers × 122 individuals × 2 thumbs per individual —6 for thumbs missing). The modal deviation was 1°, and the median deviation fell at 4°. Ninety-five per cent of the deviations were 12° or less, but it is evident that the accuracy of measurement is by no means fine. This is due to several factors. Subjective errors may be made in aligning the straight edge of the protractor with the back side of the proximal joint of the thumb, or in the exact placement of the center of the

protractor. If the subject tries hard to bend the thumb back as far as possible, fatigue rapidly sets in and the reading may decrease a few degrees. The subject will also, when asked to extend the thumb to the maximum, vary from time to time in the amount of extension. Subjects who are double-jointed at the base of the thumb (metacarpo-phalangeal joint) are particularly difficult to measure, because the degree of distal hyperextensibility is markedly affected by the position of the proximal joint at the time. This was also reported by *Harris* and *Joseph* in their study.

Harris and *Joseph* also checked the reliability of measurements by remeasuring 10 subjects on new radiographs made after an interval. The deviations ranged up to 10°, although most were under 6°. Therefore it seems that the accuracy of measurement is slightly better than in the case of our external measurements, but not enough so to warrant the extra trouble required. Two observers also independently measured the same ten radiographs three successive times. In general, as in our observations, the same observer repeated a given measurement closely, but several times there were deviations of 3° or 4°; and the two observers differed as much as 6.5° in one instance. It does appear that the considerable error involved in the external measurements makes classification of individuals uncertain unless measurements are checked by repeated trial. In a sufficiently large sample, however, the error becomes statistically unimportant. It is also worth mention that when persons rather than individual thumbs are the basis of classification, there is less error, since misclassification of a person would usually occur only when there was a coincidence of errors in measuring both the right and the left thumb. In the data of *Harris* and *Joseph* [1949] for 30 male West Africans, the mean angle of distal extensibility of the right and left thumbs was 7.35° and 10.13° greater than in the sample of male Europeans. Although the individual data are not given, there would consequently be in their sample too a greater percentage of Negroes than of Whites classified according to our criteria as having distally hyperextensible thumbs. The 31 male Indians they measured also exceeded the male Europeans, by 7.54° in the mean for the right thumb and 7.56° for the left.

Tests for Mode of Inheritance.

Data were collected on the proportions of offspring with and without distally hyperextensible thumbs (dht), in the three possible types of mating: dht × dht; dht × non-dht; and non-dht × non-dht.

These data were then tested (Table 2) by the formulae derived by *Snyder* [1934] from the *Hardy-Weinberg* principle. These formulae give the proportions, for a population breeding at random, to be expected for a single autosomal recessive gene in the matings of (1) dominant × dominant and (2) dominant × recessive. If dht is a simple recessive, then the mating dht × dht should produce only dht offspring. Actually, one exception occurred among 24 offspring in such families; but precisely such a frequency of apparent exceptions is to be expected for hyperextensibility of the thumbs, on account of its incomplete penetrance. For families in which the parents were dht × non-dht, the expected proportion of recessives is $q/(1+q)$, where q is the gene frequency of the recessive allele. The value of q may be derived from the frequency of the trait in the population, for if the trait is recessive, its frequency according to the *Hardy-Weinberg* principle will be q^2. Hence the frequency of a recessive gene is the square root of the frequency of possessors of the trait in the population. For families in which both parents are dominant, the expected frequency of recessive type offspring is $\left(\frac{q}{1+q}\right)^2$. The agreement found between the calculated values and those actually observed is excellent.

Table 2. Test for Mode of Inheritance by Sibship Formulae.

Type of parents	Number of families	Offspring dht	Offspring non-dht	Proportions among offspring dht	Proportions among offspring non-dht
dht × dht	11	30	1	.968	.032
dht × non-dht	48	37	71	.343	.657
non-dht × non-dht	132	32	281	.102	.898

A. Test for Simple Recessive Inheritance

	$q^2 = .247$			$q = .496$	Exceptions: 1/24	
$\frac{q}{1+q}$	Observed	37	71	$\left(\frac{q}{1+q}\right)^2$	Observed 32	281
	Expected	35.9	72.1		Expected 35	278
	$\chi^2 = .05$		n = 1		$\chi^2 = .29$ n = 1	
	P = .90—.80				P = .70—.50	

B. Test for Simple Dominant Inheritance.
(If dht is dominant, then q^2 = non-dht = .753)

	$q^2 = .753$			$q = .867$	Exceptions: 32/313	
$\frac{q}{1+q}$	Observed	37	71	$\left(\frac{q}{1+q}\right)^2$	Observed 1	23
	Expected	58	50		Expected 5	19
	$\chi^2 = 16.4$		n = 1		χ^2 not reliable	
	P < .01					

On the other hand, if *dht* is assumed to be a dominant gene, then the frequency of the recessive, non-dht type is 1.0–.247, or .753, and q = .867. The cross of non-dht × non-dht would in this case be expected to yield only non-dht offspring. Actually, there were 10 per cent of exceptions among 313 children of non-dht parents, a frequency too high to explain on the basis of the calculated incomplete penetrance of distal hyperextensibility of the thumbs. The results of the mating dht × non-dht also do not fit the expectancy. For the remaining type of mating, the data are too few to permit a reliable statistical test. Clearly, the hypothesis that *dht* is a dominant gene is ruled out anyway.

In Table 3 an additional test, also based on the *Hardy-Weinberg* principle, is applied. This test, described by *Dahlberg* and *Hultkrantz* [1927] and independently developed and used by *Allan* [1928, 1933] in his studies of migraine and diabetes, in certain respects has advantages over the test developed by *Snyder*. Like the latter, it requires data over two generations, but it does not require complete sibships, the index cases and their parents being sufficient. It is, of course, essential that the index cases be obtained by a true random sampling of the population. This is most easily achieved by including only one index case per family. Only if the sample is so large that sibs might be expected to occur within it on a random basis would it be permissible to include sibs, and the same might be said for other degrees of close genetic relationship. The method consists simply in comparing the *ratio* between the index cases that come, respectively, from the three types of mating: affected by affected; affected by unaffected;

Table 3. Test for Recessive Inheritance by Proportion of Total Possessors of Trait Coming from Crosses of Different Types.

	Frequency of mating	Frequency of total recessives
non-dht (Aa) × non-dht (Aa) . .	$4p^2q^2$	p^2q^2
non-dht (Aa) × dht	$2(2pq^3)$	$2pq^3$
dht × dht	q^4	q^4
$p^2q^2:2pq^3:q^4 = p^2:2pq:q^2 =$	.254+.499+.247	

	Expected	Exp. Corr.	Observed
non-dht × non-dht	23.9	25.6	26
non-dht × dht	46.9	46.8	53
dht × dht	23.2	21.6	15
		91	

$\chi^2 = 3.87$ n = 2
P = .20—.10

and unaffected by unaffected. It has often been shown that the frequency of those matings which can produce recessives is $4p^2q^2 : 4pq^3 : q^4$ (e.g., see *Boyd* [1950], pp. 435–438; *Li* [1948], pp. 17–21; *Stern* [1949], p. 167). The first type of mating produces only one-fourth offspring of the recessive type, on the average; the second, one-half; the third, all. Hence the ratio of recessives from the three classes of matings is $p^2q^2 : 2pq^3 : q^4$, obviously reducible to $p^2 : 2pq : q^2$. In other words, the proportions of recessives coming from the three types of matings in any panmictic population at equilibrium are exactly the same as the ratio in the population between the three classes, dominant homozygotes, heterozygotes, and recessive homozygotes for the character studied.

In the present case, the ratio $p^2:2pq:q^2$ has already been determined. In a total of 91 index cases forming a random sample and for whom the type of parental mating was ascertained, the expected fit to the ratio was observed ($\chi^2 = 3.87$; d.f. $= 2$; $P = .20 - .10$). This fit is not quite as good as that obtained in the first test, however, the deviation from expectation in the single class produced from the dht $\times$ dht mating being rather large ($\chi^2 = 2.9$). This is probably attributable to two factors: first, chance, because the expected size of this class is small enough to be affected considerably by sampling error; and second, the fact that this particular class will be specially affected by the incomplete penetrance of the trait. For any given dht person derived from the mating of two genotypically dht parents, there is a probability of $2 \times 0.035 = 0.07$, or 7 per cent, that one or the other parent will fail to manifest the trait on either hand. A correction of 1.62 should therefore be deducted from the expected size of the dht $\times$ dht class, and be added to the non-dht $\times$ dht class. A similar correction (amounting to 3.5 per cent) could then be applied to decrease the size of the non-dht $\times$ dht class and correspondingly increase the size of the non-dht $\times$ non-dht class. These corrections are small, but the net effect is to increase the expected size of the non-dht $\times$ non-dht class, to reduce that of the dht $\times$ dht class, and to leave the class non-dht $\times$ dht virtually unaffected. χ^2 is thus reduced to 2.84 with 2 degrees of freedom, and P is .30 to .20. The deviations are therefore attributable to chance.

The pedigrees of families with the trait clearly eliminate any possibility that it is sex-linked. On the other hand, *Li* (in press) and *Steinberg* [1952] have shown respectively that neither the *Snyder* formulae nor the $p^2:2pq:q^2$ relationship can distinguish between a

monofactorial recessive type of inheritance and a double recessive condition. Other more complicated hypotheses, such as complementary dominant genes or duplicate recessive genes, perhaps might also fit the population formulae, if the frequencies of the alleles at the interacting loci were just right. However, a simple multifactorial inheritance, dependent upon two loci without dominance and acting cumulatively (as often postulated for human skin color in crosses of caucasoid by negroid) can be ruled out. As *Rife* [1951] has shown, in such a situation the frequencies of the genotypes are as follows:

Genotype	Grade	Proportion
AABB	4	p^2u^2
AaBB, *AABb*	3	$2pu(qu+pv)$
AAbb, *AaBb*, *aaBB*	2	$p^2v^2+4pquv+q^2u^2$
Aabb, *aaBb*	1	$2qv\,(pv+qu)$
aabb	0	q^2v^2

where p and q are the frequencies of A and a respectively; and u and v are the frequencies of B and b. It follows that the ratio

$$\frac{[4]}{[0]} \text{ is the square of } \frac{[3]}{[1]}$$

This test was applied to our sample of 122 male persons, using equal ranges for the grades 0 to 4. It proved quite impossible to break the distribution into grades so that the tested relationship would fit. Instead, the ratio of the extremes was twice as great, for both right and left hands, as the ratio of classes 3 and 1, squared. This is what would be expected if the total distribution was actually bimodal but with overlapping major and minor distributions.

After this study had been completed, the trait was selected, among others, for use in comparing the gene frequencies in a religious isolate in southern Pennsylvania with those characterizing the general white population (see *Glass*, *Sacks*, *Jahn*, and *Hess* [1952]). In this small, long-isolated and highly inbred community, the frequency of distally hyperextensible thumbs is significantly lower than 24.7 per cent. However, one rather large family was found in this isolate in which 10 out of 29 members belonging to three generations have distally hyperextensible thumbs. Both parents in generation P_1 of the family are dht, and of their seven children, all were likewise dht (Fig. 5). The inheritance of the trait in this family, being clearly that of a typical recessive, confirms the conclusion previously derived regarding the mode of its inheritance.

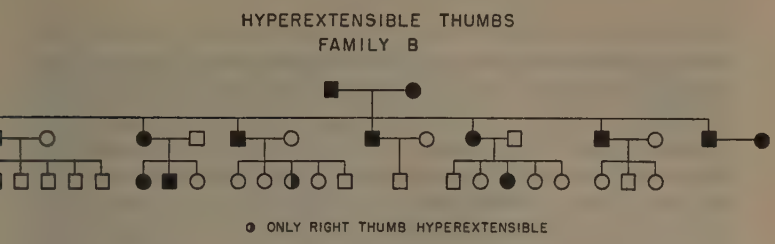


Fig. 5

Discussion.

Boyd [1950], in his excellent discussion of the need in physical anthropology for the use of genetic traits and the application of gene frequency analysis to the racial differentiation of the human species, has stated the criteria of a useful anthropological character as follows: It must be inherited according to a known mechanism, preferably simple. It must be accurately, and preferably easily, classifiable. It should be relatively stable and non-adaptive. It must be relatively common, not rare. It must vary in frequency in different populations. If it is morphological rather than physiological, and if it is a feature of parts that endure after the death of the individual, i.e., if it is skeletal, so much the better. *Boyd* comes to the conclusion that scarcely any human traits satisfy these criteria to a useable extent, except for the blood group systems, taste reaction to phenyl-thiocarbamide, the secretor gene which determines the presence or absence of blood group substances in the saliva and body fluids, and mid-digital hair. A few more traits might be added (see *Spuhler* [1950]), but the total of satisfactory ones is indeed few. One may agree emphatically with the conclusion that there is a very great need for the genetic analysis of more characters that meet these criteria and are suitable for the analysis of the genetic differences of human populations, not only for purposes of static description of present differences, but also to elucidate the dynamics of population change in genetic terms.

Hyperextensibility of the thumbs may now be added to the small group of suitable characters for this type of study, since it clearly satisfies all, or virtually all, of the criteria which have beenset forth.

It has been shown that hyperextensibility of the thumbs at the distal joint is most probably inherited as a simple Mendelian recessive character, that if so, the dominant and recessive alleles are practic-

ally equal in frequency in the white population studied, and that the frequency of the trait varies in different ethnic groups. In spite of the slightly incomplete penetrance of the recessive trait, these properties would seem to rank it as having outstanding usefulness for anthropological studies. In the accuracy with which a population may be typed for the trait it approaches the accuracy of determination of the blood types, and the ease of typing far exceeds blood group determinations and can scarcely be surpassed by any trait whatsoever. The high frequency of the trait in the two populations studied is sufficient indication that it is not subject to adverse selection, nor does it seem likely that it is subject to positive selection either. It is a fortunate triviality, a boon of Nature to the genetical anthropologist. The senior author is mildly pleased that he has thus at length realized his childhood's desire to find some kind of usefulness for his exceptional thumbs.

Summary.

Hyperextensibility of the thumb at the distal joint (dht) can be classified almost as accurately by simple external measurement as from roentgenograms. In two different samples of the U.S. white population the frequency of dht was found to be 24.7 per cent. 40 per cent of affected persons (classified as having thumbs extensible to 50° or more) have only one thumb hyperextensible. From the coincidence the penetrance may be calculated as 96.5 per cent. The trait is not noticeably affected by age or sex. It is significantly higher in North American negros (35.6 %) than in the white population.

Inheritance of the trait appears to be that of a simple Mendelian recessive, in spite of the fact that the frequency of distribution for the observed dht angles is of the continuous type. This conclusion is supported (a) by 11 pedigrees in which dht × dht matings yielded 30 dht children and only one exception, which would be expected on the basis of the calculated penetrance; (b) by a good fit to the expected ratios of affected siblings in dht × non-dht and non-dht × non-dht matings; (c) by a good fit to the expected distribution among types of matings of dht persons in a random sample of the population; and (d) by exclusion of dominant inheritance and of simple multifactorial inheritance. Inheritance through complementary or duplicate recessive genes is not excluded.

Assuming a simple recessive mode of inheritance, the frequencies of the alleles in the U.S. white population are estimated to be: *DHT*, .504; *dht*, .496. The value of a common, readily classified trait such as this in studies of human population, genetics and physical anthropology is pointed out.

Résumé.

L'hyperextensibilité de la dernière articulation du pouce (dht, c'est-à-dire « distal hyperextensibility of the thumb ») peut être déterminée à peu près avec la même exactitude en prenant des mesures sur des personnes qu'en les prenant sur des radiographies. Pour deux échantillons différents, pris dans la population blanche de l'Amérique du Nord, la fréquence de dht était de 24,7 %. 40 % des personnes atteintes (ayant des pouces extensibles jusqu'à un angle de 50° ou plus) n'avaient

que l'un des pouces hyperextensible. En se basant sur la fréquence observée on peut calculer que le taux de pénétrance est de 96,5 %. Le caractère n'est pas sensiblement influencé par l'âge ou le sexe. La fréquence plus élevée de ce caractère dans la population noire de l'Amérique du Nord (35,6 %) que dans la population blanche est significative du point de vue statistique.

Le mode de transmission du caractère semble être celui d'un caractère récessif mendélien simple, bien que la distribution de fréquence des angles de dht soit d'un type continu. Cette conclusion se base *a*) sur 11 arbres généalogiques dans lesquels des unions dht × dht donnaient 30 enfants atteints de dht et seulement une exception, comme on pouvait s'y attendre par suite de la pénétrance calculée; *b*) sur une concordance satisfaisante avec les taux prévus des descendants atteints dans des unions entre des personnes dht–non dht, et entre des personnes non dht–non dht; *c*) sur une concordance satisfaisante avec la distribution prévue d'enfants atteints provenant d'unions, prises au hasard dans la population, entre des personnes dont l'une est atteinte de dht; et *d*) sur l'exclusion de transmission dominante et de transmission plurifactorielle simple. L'hérédité par « complementary or duplicate recessive genes » ne peut pas être exclue.

Si l'on admet que le mode de transmission est simple récessif, la fréquence des allèles dans la population blanche de l'Amérique du Nord peut être évaluée à: *DHT*, .504; *dht*, .496. Du point de vue génétique et anthropologique on souligne la valeur d'un tel caractère, à la fois fréquent et facile à classifier, pour des études de populations humaines.

Zusammenfassung.

Distale Hyperextensibilität des Daumens (dht = distal hyperextensibility of the thumb) kann nahezu gleich exakt durch einfache externe Messung wie auch durch Röntgenogramm klassifiziert werden. In zwei verschiedenen Auswahlen unter der weißen U.S.-Bevölkerung wurde für die dht Frequenz der Wert von 24,7 % gefunden. 40 % der behafteten Personen (mit einer Daumenextensibilität von 50° und mehr) haben dht nur bei einem Daumen. Aus der Koinzidenz kann man eine Penetranz mit 96,5 % annehmen. Das Merkmal ist nicht wesentlich durch Alter und Geschlecht beeinflußt. Es kommt sichtlich häufiger bei nord-amerikanischen Negern (35,6 %) als bei der weißen Bevölkerung vor.

Die Vererbbarkeit des Merkmals scheint einfach rezessiv zu sein, trotz der Tatsache, daß die Verteilungskurve der beobachteten dht Winkel kontinuierlich ist. Diese Feststellung wird unterstrichen durch: *a*) 11 Vererbungstafeln, in den dht × dht Kreuzungen 30 dht Kinder ergaben mit nur einer Ausnahme, die man auf Grund der einberechneten Penetranz erwarten durfte; *b*) eine gute Übereinstimmung mit der erwarteten Proportion behafteter Geschwister in dht × nicht-dht und nicht-dht × nicht-dht Kreuzungen; *c*) eine gute Übereinstimmung mit der erwarteten Verteilung unter den Typen bei Kreuzungen von dht-Personen in einer zufälligen Auswahl unter der Bevölkerung und *d*) durch den Ausschluß dominanter Vererbung und einfacher Polymerie. Vererbung durch «complementary or duplicate recessive genes» ist nicht ausgeschlossen.

In der Annahme eines einfachen rezessiven Erbganges schätzt man die Allellfrequenzen in der weißen U.S.-Bevölkerung auf *DHT* 0.504; *dht* 0.496. Der Wert eines gewöhnlichen, bereits klassifizierten Erbmerkmales wie dieses in Studien menschlicher Populationen aus genetischen und physisch-anthropologischen Gesichtspunkten wird herausgestellt.

REFERENCES

Allan, Wm.: Arch. intern. Med. *42*, 590–599, 1928; *6*, 1272–1274, 1932. – *Boyd, W. C.:* Genetics and the Races of Man. Little, Brown and Co., Boston 1950. – *Dahlberg, G.* and *J. V. Hultkrantz:* Arch. Rass. Ges. Biol. *19*, 129–165, 1927. – *Glass, B., M. S. Sacks, E. F. Jahn* and *C. Hess:* Amer. Nat. *86*, 145–159, 1952. – *Harris, H.* and *J. Joseph:* J. Bone Jt. Surg. *31*B, 547–559, 1949. – *Li, C. C.:* An Introduction to Population Genetics, National Peking University Press, Peiping 1948; Amer. J. hum. Genet. (in press). – *Rife, D. C.:* J. Hered. *42*, 162, 1951. – *Snyder, L. H.:* Genetics *19*, 1–17, 1934. – *Spuhler, J. N.:* Cold Spr. Harb. Symp. quant. Biol. *15*, 175–189, 1950. – *Steinberg, A. G.:* Amer. J. hum. Genet. *4*, 373–375, 1952. – *Stern, C.:* Principles of Human Genetics, W. H. Freemann, San Francisco 1949. – *Whitney, D. D.:* Family Treasures, J. Cattell Press, Lancaster 1942. – *Whitney, L. F.:* J. Hered. *23*, 425–426, 1932.

---

von Hofsten, N.: Acta genet. *4*, 206–221, 1953

ON THE THEORETICAL EFFECT OF MUTATION

By NILS von HOFSTEN

The effect of mutation proceeding at a steady rate is easily calculated, but an analysis of the process is perhaps not superfluous; it gives results which, though rather self-evident, merit attention or, at any rate, have a theoretical interest.

If a character is due to a single gene, the increase of the character-carriers through mutation is, in cases of typical dominance-recessivity, a simple function of that of the gene. What I intend to discuss especially is the effect of mutation affecting two different genes, above all in cases of dimerism where a character is due to two genes.

It is presumed in all following calculations and deductions that mutation is proceeding undisturbed by selection, back-mutation and mutation of the mutated genes. When a large number of generations is considered, this presumption is very unnatural. This does not affect the principles which are of importance for the present purpose, and the comparison between different cases is facilitated if such disturbing factors are left out of account. In science it is often essential to select one single factor from several involved in a process. A purely theoretical investigation has the advantage that this is always possible. Even when applied to the point of absurdity, this method may give valuable insight into a process.

The same argument holds good for the presumption of a very large or exceedingly large number of generations. Such a presumption may elucidate changes too insignificant to appear—or to be illustrated in a diagram—in a moderate number of generations. Furthermore, what constitutes a moderate and what constitutes an excessively large number is, of course, dependent on the speed of reproduction. In higher animals and plants 10000 generations usually require as many years or more (in man about 300000 years), in Drosophila (theoretically) only 400 years. In bacteria 100000 generations may be accomplished in 4 years.

For the present purpose there is no need to discuss the possible or probable relations between the rate of propagation, i.e. the time factor, and the mutation frequencies, or the question of differences between different groups of organisms.

Even in a very long series of generations mutations of rates known to occur (mostly 1 : 20 000 to 1 : 1 000 000 and less) give, if selection is disregarded, only a slight rise of the gene frequency. It is, on this account, advisable to consider the result of stronger increases of the gene frequency (more about this below).

Only recessive gene mutations, i.e. changes of dominant genes to recessive (D to R), are considered.

The question of interaction between mutation and selection will not be discussed. Some remarks on this subject, and the equilibrium which according to the important deductions of *Haldane* will be established, are given in a previous paper on the theoretical effect of selection[1].

The effect of mutation on the frequency of a gene.

The basis of investigations of this kind is always, of course, the effect of mutation on the frequency of a single gene.

The frequency of a mutated gene is calculated by means of a formula given by *Dahlberg*[2]. If the mutation rate is μ, the gene frequency in the nth generation is $1-(1-\mu)^n$. *Dahlberg* states that in a low mutation rate the gene composition will at first not be noticeably changed; "only after a long time does the effect begin to be noticeable". This is correct, of course, but does not signify that the

[1] *N. v. Hofsten*, The genetic effect of negative selection in man. Hereditas *37*, 1951.

[2] *G. Dahlberg*, Mathematical methods for population genetics, 1947 (p. 38 seq.). Cf. also *L. Hogben*, An introduction to mathematical genetics, 1946, p. 190 seq.

effect is weaker at the beginning and gradually gets stronger. It is easily understood that quite the opposite must be the case: the effect must be stronger at first when the mutating genes are very numerous and become weaker as the supply of old, non-mutated genes diminishes.

For a low mutation frequency and a more or less limited number of generations, it is scarcely possible to get a clear insight into the process. Even at the high mutation rates of 0.0001 and 0.001 (the latter extremely high) the increase in the gene frequency follows in a large number of generations a seemingly straight line. There are two ways to obtain a true picture of what really occurs: to analyse the effect of a very exaggerated mutation frequency, perhaps some thousand times stronger than what commonly occurs in natural conditions, or to consider an extremely large number of generations. For very high but less exaggerated mutation rates both methods may be combined.

In the following, the word "mutation" (in quotation marks) means a process analogous to mutation but presumed to occur in a much higher frequency. It should be observed, however, that surprisingly high mutation rates seem to occur quite exceptionally and that the spontaneous mutation frequency, as is well known, may be increased very strongly by means of agents inducing mutation. Thus, even mutation rates of 0.001 to 0.01 are not wholly fictitious.

If we presume an increase of 0.1, i.e. a preposterously high "mutation" frequency, it becomes at once apparent that the increase of the mutated genes is at first strong and then gradually diminishes (fig. 1). In the 6th to 7th generation the new and the old gene have the same frequency. After 30 generations not much remains of the mutating dominant genes, and the frequency of the new ones rises slowly. In Gen. L the frequency will be 99.5, in Gen. LX 99.8 per cent. The increase is in every generation 0.9 of the increase in the previous generation. In other words: if in Gen. n the increase is c, it is $(1—\mu)\ c$ in Gen. $n+1$. Mathematically, a remnant of the supply of the old dominant genes will be left in the population, finally one single gene (and still less), but this can be said to ressemble the problem of Achilles and the tortoise.

The same general rules must be valid for lower "mutation" frequencies and for real mutations. The frequency curves are always identical and may all be derived from an arbitrary curve of this kind, for instance the curve for $\mu = 0.1$ (fig. 1), through change of the

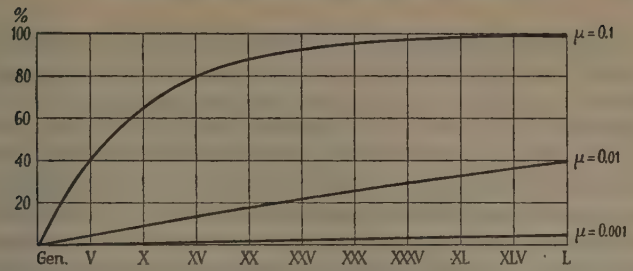


Fig. 1. Effect of mutation. Increase of the frequency of a new recessive gene. Three different mutation rates (exaggerated), 0.1, 0.01, 0.001. 50 generations.

scale of the horizontal n-axis, i.e. extension of this axis when going from a higher to a lower mutation rate. There is, however, no exact correspondence between the generations. If the rates $\mu_1 = 0.1$ and $\mu_2 = 0.01$ are compared, 1 generation in the former case corresponds to 10.5 generations in the latter case. Thus, the percental frequencies 10, 19, 27.1, 34.39, etc., which are at $\mu = 0.1$ reached in Gen. I, II, III, IV, etc., are at $\mu = 0.01$ attained in Gen. "10.5", XXI, "31.5", XLII, etc. When the generations are compared, $\mu = 0.01$ gives in Gen. X, XX, XXX and XL the frequencies 9.56, 18.21, 26.03, 33.10 per cent[1].

On practical grounds it is convenient that curves for different mutation rates and numbers of generations are drawn in a scale which, so to speak, is based on the relation between the mutation rates, i.e. where in the examples given above the same distance covers, at $\mu = 0.1$ one, at $\mu = 0.01$ ten, at $\mu = 0.001$ hundred generations, etc. (cf. figs. 1, 2 and the diagrams of homozygote frequencies). The curves then do not coincide but the difference is not great. If a gene R at $\mu = 0.1$ attains the frequency r_x in Gen. n, it attains in lower mutation rates the same frequency in the neighbourhood of but succeeding, at $\mu = 0.01$: Gen. $10 \cdot n$, at $\mu = 0.001$: Gen. $100 \cdot n$, at μ 0.0001: Gen. $1000 \cdot n$, etc.

What is of interest in this connection is above all a concrete comparison between the effect of mutation at different rates.

---

[1] Dr. *B. Kjellberg*, with whom I had the opportunity of discussing this matter, has made the following observation: To obtain from $\mu_1 = 0.1$ the value of r in Gen. n at $\mu_2 = 0.01$, $1-\mu_1$ has to be raised to the power of $k \cdot n$ where $k = \frac{\log. 0.99}{\log. 0.9}$. Hence, $r_2 = 1-(1-\mu_2)^n = 1-(1-\mu_1)^{k \cdot n}$.

A "mutation" rate of 0.01 gives a curve which in the scale of the diagram fig. 1 is feebly curved. The increase is in every generation (1—0.01 =) 0.99 of that in the foregoing generation. After 50 generations the frequency of the mutated gene is 39.5 per cent; it will be almost 99.5 per cent after 500 and after still 100 generations have risen very slowly to 99.8 (99.82) per cent. At a mutation rate of 0.001 the increase is in every generation reduced to (1—μ =) 0.999, and the frequency curve follows in the scale of fig. 1 a seemingly straight line (in reality imperceptibly bowed). The frequencies 99.5 and 99.8 per cent will be reached after more than 5000 and 6000 generations respectively.

If a very great number of generations is considered (50000 in fig. 2), the frequency curve for $\mu = 0.001$, which attains 99.5 per cent in 5000 generations, of course in the scale of this diagram rises very steeply. The curve for $\mu = 0.0001$ is curved and ressembles that for $\mu = 0.1$ in fig. 1 (fig. 2). The reduction in the increase is basically the

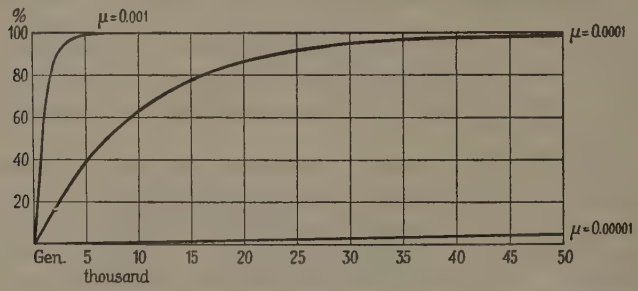


Fig. 2. As fig. 1, mutation rates 0.001, 0.0001, 0.00001. 50 000 generations.

same as in higher mutation rates (0.9999). A frequency of 99.8 (more exactly 99.75) per cent is reached in 60000 generations.

Finally we presume a mutation rate of an order of magnitude which may be said to be typical for many well-known mutations. In a mutation rate of 0.00001 (1:100 000) the gene frequency rises exceedingly slowly and the theoretical reduction of the increase from one generation to another is almost imperceptible (0.99999). The frequency rises in 100 generations to 0.01, in 1000 to 0.1, in 10 000 to almost 1 and in 50 000 generations to a little less than 5 per cent. The curve is in the scale of the diagram fig. 2 a practically straight line. An increase to 50 per cent would require more than 600 000

generations, an increase to 99.8 per cent more than 6 million generations.

Monomerism. Increase of the recessive homozygotes.

It is hardly necessary to mention that in every generation the frequency of the recessive homozygotes is the square of the frequency of the recessive genes, but this fact has a consequence that should be observed. It is presumed in the following that mutation *D* to *R* sets in in a population *DD* (initially no recessive alleles). It is sufficient to consider the effect of a "mutation" of 0.1.

As shown by the diagram fig. 3 the r^2-curve, contrary to the gradually straightened *r*-curve, begins with a bending downwards; it has, roughly speaking, the shape of an irregular, very extended S. This difference, perhaps unexpected to a superficial observer, is not difficult to understand. When (in the last part of the both curves) *r* approaches 1 (100 per cent), the transition from *r* to r^2 signifies a more and more insignificant percental change. On the other hand, when in the beginning of the curve *r* has risen but little above zero, r^2 is only a small percentage of *r*. The reverse occurs, i.e. the inflection point of the curve lies where *r* has reached the frequency 0,5. The diagram shows that in this case (μ 0,1) this happens between Gen. VI and Gen. VII[1].

The relation between gene frequency and homozygote frequency, though mathematically extremely simple, has consequences which

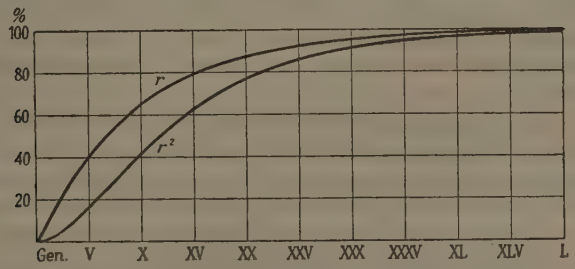


Fig. 3. Mutation 0.1. Lower curve, monomeric recessive homozygotes (*RR*). Upper curve, the new gene (*R*). 50 generations.

[1] I am indebted to Dr. *B. Kjellberg* for an instructive discussion of the inflection of the r^2-curve. He has remarked that the point of inflection is obtained in Gen. n_i when n_i satisfies the equation $(1-\mu)^{n_i} = 1/2$; $n_i = \frac{\log. 2}{\log. \frac{1}{1-\mu}}$.

are of considerable interest from a genetic point of view. The frequency of the character-carriers, the recessive homozygotes, increases in the first generations, as is clearly shown by the curves in fig. 3, comparatively less than the frequency of the genes. In the first generation the difference, as was emphasized before, is very great (r 0.1, r^2 0.01). As a consequence of the low relative frequency of the recessive homozygotes in the first generation, their increase from one generation to the succeeding one is very strong in the next few generations; they so to speak recover a part of the delay. Their increase remains for a long time relatively stronger than the increase of the genes. While the increase of the genes is constantly 0.9 of the increase in the preceding generation (p. 208), the corresponding figures for the homozygotes are: Gen. I–II: 3.61, II–III: 1.43, III–IV: 1.20, IV–V: 1.10, V–VI: 1.05, etc.; the figure is more and more slowly going down towards 0.9 (Gen. IX–X: 0.96, XXII–XXIII: 0.91).

What has been said above about the effect of lower mutation rates on the increase of the genes is of course, *mutatis mutandis*, applicable also to the increase of the homozygotes.

Dimerism. Initial population $D_1D_1D_2D_2$ (no recessive alleles). Mutation D_1 to R_1 and D_2 to R_2.

We assume that D_1 and D_2 have the same frequency ($d_1 = d_2$) and that in both genes the mutation rate is the same; R_1 and R_2 consequently have the same frequency in all generations. (This is the simplest case; other cases will not be analysed in this connection.) The frequency of the double homozygotes in the nth generation is the frequency of any of the two genes raised to the fourth power: $[1-(1-\mu)^n]^4$.

In the diagram given below (fig. 4) the upper curve shows the frequency of the gene combination R_1R_2, which—because of the equal frequency of the genes—also is the frequency of R_1R_1 and of R_2R_2. The lower curve gives the frequency ($r_1^2 r_2^2$) of the double recessive homozygotes $R_1R_1R_2R_2$. The diagram thus illustrates the different result in the two processes presumed: mutation of one gene (D to R) and mutation of two genes (D_1 to R_1, D_2 to R_2) (under the conditions named above, equal frequency of the genes, the same mutation rate). It is natural to presume that the two genes influence the same character, and the comparisons have more interest if this is presumed, but the result is, of course, independent of the phenotypical effect of the genes.

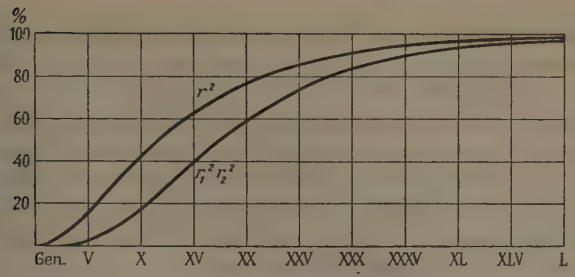


Fig. 4. Mutation 0.1. Lower curve, dimerism (new recessive genes R_1 and R_2), $r_1^2 r_2^2$ (= r_1^4 and r_2^4). Upper curve, monomerism, r^2 (= $r_1 r_2$). 50 generations.

It is a matter of course that the effect of mutation is weaker when two mutating genes are involved. As the initial frequency of the recessive genes and homozygotes is zero, in dimerism as well as in monomerism, and the final result—if the mutation is continued long enough—will be a frequency of 100 per cent, the curves must diverge at the beginning and sooner or later converge again and ultimately meet. From what has been said about the increase of monomeric recessive homozygotes it is obvious that the frequency curves of double homozygotes also have the extended S-form discussed above (p. 211).

If a "mutation" rate of 0.1 is presumed (fig. 4), the frequency of $R_1R_1R_2R_2$, from 1 per cent of the frequency of *RR* in Gen. I, gradually rises to 100 per cent (Gen. II: 3.6, III: 7.3 .... X: 42.4, etc.). The distance between the curves, i.e. the difference between the frequencies, attains its maximum in Gen. XII (frequencies 51.49 and 26.5 per cent respectively). In the 50th generation *RR* (the upper curve) have almost entirely replaced the dominant genotypes (*RR* frequency 98,97 per cent) and $R_1R_1R_2R_2$ have nearly the same frequency (97.95 per cent). The few remaining dominant genes are eliminated extremely slowly; after a further 50 generations there are 99.47 per cent *RR* and 98.94 per cent $R_1R_1R_2R_2$.

In lower mutation rates the processes are principally the same but the effect is of course weaker and, consequently, postponed. If we pass directly to a mutation rate of 0.0001, which is high but occurs, curves for 50 generations would have no meaning, for the frequency of *RR* increases only to 0.0025 per cent, while that of $R_1R_1R_2R_2$ remains practically unaltered (0.000625 in 1 million). If the effect is calculated for some ten thousand generations, it appears

that the effect is quite analogous to that in the former case (μ 0.1); from causes explained above (p. 208 seq.) the curves (r^2 and $r_1^2\ r_2^2$) are, in an appropriate scale, identical with those at μ 0.1, though the generations do not correspond. Fig. 4 therefore illustrates fairly well the increase in the mutation rate 0.0001, if the numbers of the generations are multiplied by 1000. Mutation must go on for more than 50 000 generations before the curves definitely approach each other and a frequency of 100 per cent (Gen. 50 000 : 98.65 and 97.32 per cent, 60000: 99.5 and 99.0 per cent).

Concerning the effect of mutation in the more typical rate of 0.00001 two items are sufficient. In 100 generations the frequency of *RR* will be only 1 in 100 millions (10^{-8}), which means that in panmixia in a fairly large population they are practically non-existent. The chance of the formation of a double homozygote $R_1R_1R_2R_2$ is so utterly insignificant (10^{-16}) that it must be said to be non-existent. After 1000 generations (30000 years in man!) the chance would (in panmixia, etc.) be for *RR* 1:1 million and for $R_1R_1R_2R_2$ still practically nil ($10^{-11} = 1:1000$ milliards). It would require 6 million generations to give the new recessive type an almost complete supremacy in the population (*RR* 99.47, $R_1R_1R_2R_2$ 98.95 per cent).

Dimerism. Initial population containing a certain amount of double recessive homozygotes ($R_1R_1R_2R_2$). Comparison with the effect in monomerism.

If a character is due to two genes, mutation affecting these genes or one of them has, of course, another effect than in monomerism.

Let us assume that a character, not influenced by environment, is due to two recessive genes and that the mutation, where both genes are concerned, has the same rate. In all examples discussed in the following the initial frequency of the recessive homozygotes is assumed to be 1 per cent.

1. The two recessive genes (and their dominant alleles) have the same frequency.

If r is the frequency of each of the recessive genes, μ the mutation rate, and the first generation changed by the mutation is taken as the first, the frequency of the gene in the nth generation is

$$1 - (1-r)\ (1-\mu)^n$$

and the frequency of the character-carriers (double recessive homozygotes) is

$$[1-(1-r)(1-\mu)^n]^4$$

We now presume an initial frequency of the character-carriers of 0.01 (1 per cent). This frequency implies that the two recessive genes have each a frequency of 0.31622666... To get a clear conception of the effect, an impossibly high "mutation" rate of 0.1 is presumed.

The effect of this "mutation" in 25 generations is illustrated by the lower continuous curve in the diagram fig. 5; the upper continuous curve shows, as in the previous examples, the corresponding effect on the frequency of a monomeric recessive character. It is easily understood that the curves, as when the starting point is zero, have a diverging-converging course (the distance attains its maximum in Gen. X). From causes explained before (p. 211, cf. p. 213) it is also obvious that the $r_1^2r_2^2$-curve has the same characteristic shape as the r^2-curve. The frequency is in dimerism raised relatively strongly in the first generation, to 60.6 per cent of the frequency in monomerism. This percentage sinks in the two following generations (to 54.0 and 53.5 respectively) and then slowly rises. In Gen. XXV the frequencies are 87.5 and 81.8 (93.5 per cent of 87.5) respectively. In Gen. LX the recessive homozygotes will have a frequency of about 99.3 and 98.6 per cent respectively.

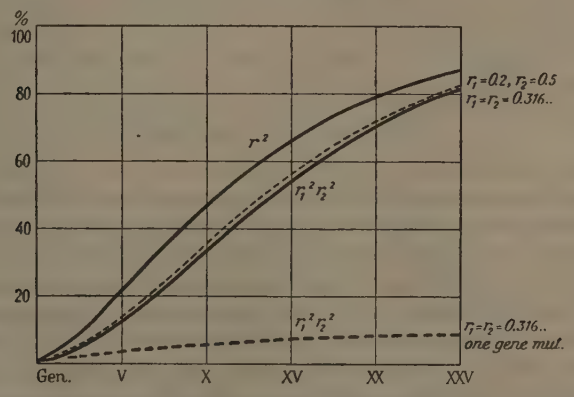


Fig. 5. Recessive homozygotes initially 0.01. Mutation 0.1. Lower continuous curve, dimerism, $r_1^2r_2^2$. Upper continuous curve, monomerism, r^2 ($= r_1r_2$ in dimerism). Broken lines: Upper curve, $r_1 = 0.2$, $r_2 = 0.5$, both genes mutated. Lower curve, r_1 and $r_2 = 0.3162266..$, One gene mutated. 25 generations.

The curves for lower mutation rates are, in the sense explained above (generations not corresponding, cf. p. 209), identical. The curves in fig. 5 could therefore, if drawn for a greater number of generations, replace special curves. Nevertheless I think that two diagrams and a few remarks, though rather monotonous, will give a clearer insight into the examples and thus be justified.

At a "mutation" rate of 0.01 the curves show of course the same characteristics as in fig. 5, though all changes are postponed (10 generations = a little less than 1 at $\mu = 0.1$) (fig. 6). The frequency is at first raised relatively high, then the increase is diminished until Gen. XXX (Gen. I 91.7 per cent of the frequency in monomerism, Gen. II 85.1... Gen XXX 53.1 per cent); from the next generation it slowly rises. The distance between the curves has its maximum in the 100th generation. After 500 generations the frequencies of RR and $R_1R_1R_2R_2$ are 98.8 and 98.2 per cent and after a further 100 generations they will almost be the only individuals left and have nearly the same frequency (99.7 and 99.4 per cent).

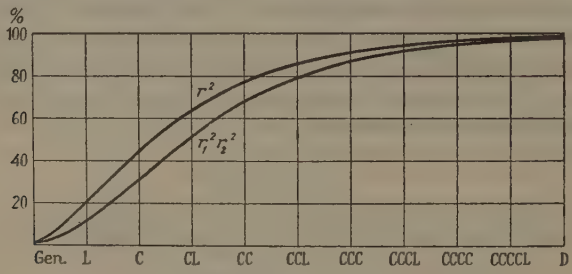


Fig. 6. As fig. 5 but mutation 0.01. 500 generations.

At a mutation rate of 0.0001 (which may occur) the curves have the same shape, when drawn for an extremely large number of generations (fig. 7). One generation at $\mu = 0.1$ corresponds to about 1000 at $\mu = 0.0001$. The increase from one generation to the next one is of course extremely small; in the first generation the frequency is raised in monomerism from 1 to 1.0018 and in dimerism to 1.00085 per cent. The increase thus being at first, in monomerism as well as in dimerism, very little more than zero, the difference in frequency is in several generations very small; in Gen. I the frequency in dimerism is 99.905 per cent of the frequency in monomerism. The percentage sinks slowly (91.7 in Gen. 100), rises a little, sinks to 53,4 in Gen. 3000

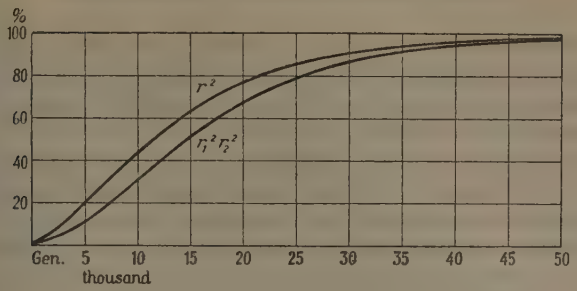


Fig. 7. As fig. 5 but mutation 0.0001. 50 000 generations.

and from then rises to the final value 100. A frequency of almost 100 per cent is attained after more than 60000 generations. — At a mutation rate of 0.00001 the same result would require more than 6 million generations.

2. R_1 *and* R_2 *have different frequencies.*

If the genes have initially the frequencies r_1 and r_2 and μ is the mutation rate, the recessive homozygotes have in the nth generation the frequency

$$[1-(1-r_1)(1-\mu)^n]^2 \cdot [1-(1-r_2)(1-\mu)^n]^2.$$

There is, of course, an almost infinite number of combinations between different frequencies of two genes R_1 and R_2. We chose the gene combination 0.8 D_1:0.2 R_1:0.5 D_2:0.5 R_2, which gives, as in the previous case, a frequency of the recessive homozygotes $R_1R_1R_2R_2$ of 0.01 (1 per cent), and presume a "mutation" rate (impossibly high) of 0.1.

Because of the lower frequency of one of the recessive genes the effect of mutation is stronger than in equal frequency of R_1 and R_2 but, as shown by fig. 5 (upper broken curve) and fig. 8, the difference is, in this example, not great. The surplus is, in per cent of the frequency attained in the case of $r_1 = r_2$, in Gen. I 8.4, rises to 10.9 in Gen. III and then steadily falls to 1 per cent in Gen. XXV and finally to zero. In Gen. XIV the recessive homozygotes have increased to about half of the population (52.4 per cent), in Gen. XXV to 82.6 per cent. A frequency of 99.5 per cent is attained in 60 generations. The difference between r_1 and r_2 is gradually eliminated (Gen. 0:0.20 and 0.50; Gen. V: 0.53 and 0.70; Gen. X: 0.72 and 0.83; Gen. XV: 0.84 and 0.90; Gen. XX: 0.90 and 0.94).

From the analysis of dimerism with equal frequency of R_1 and R_2 it is easily understood that lower mutation rates give essentially the same result though a much larger number of generations is required to get corresponding frequencies.

The greater the initial difference of frequency between the genes, the stronger is the effect of the mutation, and the more it approaches the effect obtained in monomeric recessivity (the same initial frequency of the recessive homozygotes presumed). The reason is very simple: when a recessive gene is frequent, its dominant allele is rare, and the rarer it is the less is the effect of its mutation, i.e. the conditions approach those where the other recessive gene alone is present. If, for instance, the initial frequencies of R_1 and R_2 are 0.11270166... and 0.88729833..., which gives a $R_1R_1R_2R_2$ frequency of 1 per cent, a mutation rate of 0.1 gives the following frequencies (in brackets the corresponding frequencies in monomeric recessivity): Gen. I 3.276 (3.61), Gen. V 19.75 (21.95), Gen. X 44.02 (47.06), Gen. XXV 86.65 (87.50) per cent (fig. 8).

3. One gene only is mutating.

When only one of the recessive genes is increased by mutation and the other maintains its frequency, the effect of the mutation is,

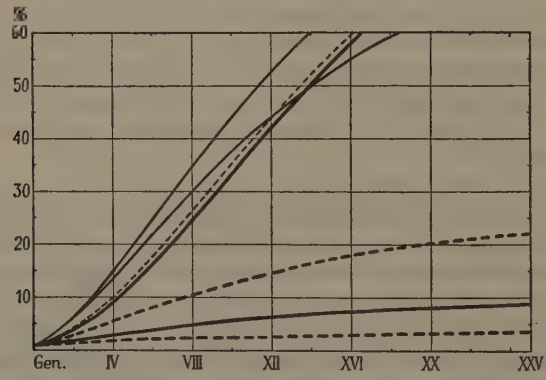


Fig. 8. Dimerism. Recessive homozygotes initially 0.01. Mutation 0.1. Thick continuous lines: r_1 and r_2 0.3162266..; upper curve, mutation of both genes; lower curve, mutation of only one gene. Broken lines: r_1 0.2, r_2 0.5; upper curve (fine line), mutation of both genes; middle curve, mutation D_1 to R_1; lower curve, mutation D_2 to R_2. Fine continuous lines: r_1 0.11270166.., r_2 0.88729833..; upper curve, mutation of both genes; lower curve, mutation D_1 to R_1. 25 generations.

when the frequency of the genes is the same or not very different, more or less heavily reduced. Let us consider a simple case: initially r_1 and r_2 both 0.3162266... and $r_1^2 r_2^2$ 0.01, R_1 (or R_2) increased through "mutation" at a rate of 0.1. In the first generation the homozygote frequency (1.48 per cent) lies not very much below the frequency in mutation of both genes (2.19 per cent) but it rises increasingly slowly; the difference rapidly increases and the curves diverge very strongly (fig. 8, cf. fig. 5). The frequencies in the two cases are in Gen. V 3.56 and 12.64, in Gen. X 5.80 and 33.64, in Gen. XX 8.41 and 70.67, in Gen. XXV 8.99 and 81.76, in Gen. LX 9.975 and 99.51 per cent. A frequency of 10 per cent is the limit, after which no further increase occurs. This limit is of course due to the fact that the frequency is dependent on the non-mutated gene only, when the mutated gene has totally replaced its dominant allele (r_2 0.316..., r_2^2 0.1).

When the two genes have different frequencies, the result depends on which of the genes is mutated. The effect is stronger when the frequency of the rarer gene is increased. This is due to the constant influence of the other gene, whose frequency remains unaltered. If the gene relation is r_1 0.2 : r_2 0.5 ($r_1^2 r_2^2$ 0.01) and only R_1 is increased by "mutation" at the rate 0.1, the increase of the double homozygotes $R_1R_1R_2R_2$, though much less than if both genes are mutated, is stronger than in the case of r_1 and r_2 0.3162266... and mutation of one gene (fig. 8). In Gen. XXV the frequency is 22.26 per cent. The limit 25 per cent (frequency of $r_2{}^2$) will soon be attained (Gen. XXX 24.8 per cent). If, on the contrary, only R_2 is increased, the increase is much less than when the genes have the same frequency (fig. 8). In Gen. XXV the frequency is 3.75 per cent, and the limit 4 per cent is nearly attained in Gen. XXX (3.98 per cent). It is interesting to note that the frequency is doubled (2 per cent) already in the fourth generation.

The more one of the genes prevails, the stronger is the effect of the increase of the other, rarer gene. If the gene frequencies are r_1 0.11270166... and r_2 0.88729833... ($r_1^2 r_2^2$ 0.01 as in all examples discussed) and only the rare gene r_1 is increased by mutation (0.1), the effect on the homozygote frequency is rather marked and in several generations not much less than if both genes are mutated. Gradually the curve owing to the influence of the non-mutated gene becomes more and more marked and crosses two of the other curves in fig. 8. If, on the other hand, the frequent gene R_2 is

increased, the effect, as is easily understood, is extremely small; the homozygote frequency is in 25 generations increased only from 1 to 1.25 per cent, which lies near to the upper limit (1.27 per cent). A curve would in the scale of the diagram be a straight and almost horizontal line.

I have not paid attention to the hypothetic case of two genes mutating at different rates. There is, of course, a series of transitions between mutations at equal rates and mutation of one gene only. The greater the difference between the rates, the more the effect approaches the result attained when only one of the genes mutates.

Summary.

A purely theoretical analysis is given of the effect of mutation proceeding at a steady rate and undisturbed by selection, etc. In several cases an exceedingly large number of generations is presumed, and also a very exaggerated mutation rate. Such more or less unnatural presumptions do not affect the principles, and they give a clear insight into the process. Only recessive mutations (*D* to *R*) are considered.

The basis of the deductions is the effect of mutation on the frequency of a single gene (figs. 1, 2). The frequency of the recessive homozygotes is easily calculated; the frequency curve has an interesting form (fig. 3). The effect in cases of dimerism is discussed and compared with the effect in monomerism (fig. 4). Such comparisons are of special interest when the inital population contains a certain number of recessive homozygotes (figs. 5–8). In these cases the effect of mutation is calculated under various presumptions: equal frequency of the genes; different frequencies of the genes; mutation of both genes; mutation of one of the genes.

Résumé.

L'auteur fait une analyse purement théorique de l'effet de mutations qui ont une fréquence constante et dont le résultat n'est pas influencé par sélection etc. Dans plusieurs cas on suppose un nombre de générations extrêmement élevé ainsi qu'une fréquence de mutations très exagérée. De telles suppositions plus ou moins artificielles n'influencent pas les principes et donnent une idée nette du processus. On considère seulement les mutations récessives (mutations de *D* en *R*).

L'effet de la mutation sur la fréquence d'un gène unique fournit la base des déductions (fig. 1, 2). La fréquence des homozygotes récessifs est facile à calculer; la courbe de fréquence a une forme intéressante (fig. 3). L'auteur discute l'effet de mutation dans les cas de dimérie et le compare avec le même effet dans les cas de monomérie (fig. 4). De telles comparaisons offrent un intérêt particulier quand la population initiale contient un certain nombre d'homozygotes récessifs (fig. 5–8). Dans ces cas l'effet de la mutation est calculé par rapport à des suppositions différentes: fréquence égale des gènes; fréquence différente des gènes; mutation des deux gènes; mutation d'un des gènes.

Zusammenfassung.

Der Verfasser führt eine rein theoretische Analyse des Effektes der Mutation bei konstanter Mutationsfrequenz und ohne Beeinflussung durch Selektion usw. durch. In mehreren Fällen wird eine außerordentlich große Anzahl von Generationen angenommen sowie eine sehr übertriebene Mutationsfrequenz. Derartige mehr oder weniger unnatürliche Annahmen beeinflussen die Prinzipien nicht und gewähren einen klaren Einblick in den Vorgang. Nur rezessive Mutationen (*D* in *R*) kommen in Betracht.

Den Schlußsätzen liegt die Wirkung der Mutation auf die Frequenz eines einzelnen Genes zugrunde (Abb. 1 und 2). Die Frequenz der rezessiven Homozygoten ist leicht zu berechnen; die Frequenzkurve hat eine interessante Form (Abb. 3). Die Wirkung in Fällen von Dimerie wird erörtert und mit derjenigen in Monomerie verglichen (Abb 4). Solche Vergleiche besitzen ein spezielles Interesse, wenn die ursprüngliche Population eine gewisse Anzahl rezessiver Homozygoten enthält (Abb. 5–8). In diesen Fällen wird der Mutationseffekt unter verschiedenen Voraussetzungen berechnet: gleiche Frequenz der Gene; ungleiche Frequenz der Gene; Mutation beider Gene; Mutation von einem der Gene.

Johansson, I.: Acta genet. *4*, 221–231, 1953

Institute of Animal Breeding, Royal Agricultural College of Sweden, Uppsala

THE MANIFESTATION AND HERITABILITY OF QUANTITATIVE CHARACTERS IN DAIRY CATTLE UNDER DIFFERENT ENVIRONMENTAL CONDITIONS

By IVAR JOHANSSON

Most of the economically important characters in farm animals are quantitative, showing a continuous variation. When the variates are thrown into classes of equal intervals, the frequency distribution approximates the normal curve of error. In regard to these characters a strictly Mendelian analysis is not practicable because each character depends, in most cases, on a large number of genes whose individual effect cannot be isolated, and the phenotypic manifestation is subject to considerable modifications by environmental

influences. Growth rate, body size at maturity, milk yield and butterfat content of the milk may be mentioned as examples of such characters. For the planning of breed improvement it is, fortunately, rather immaterial whether we know *how* these characters are inherited or not, but it is very important to know how large a part of the total variation is controlled by heredity, i.e. their *heritability*.

Lush [1949] has defined heritability, in the narrow sense of the term, as the fraction of the total observed variance which is caused by the average (i.e. the additive) effect of the genes. The total variance (σ_p^2) may be divided into different portions according to the genetic and environmental causes of phenotypic variation, expressed as deviations from the population average, as follows:

$$\sigma_p^2 = \sigma_g^2 + \sigma_d^2 + \sigma_i^2 + \sigma_e^2 + \sigma_j^2$$

where g denotes additive genetic deviations, d dominance deviations, i epistatic deviations or gene interactions, e deviations due to environment, and j represents the non-linear interactions between heredity and environment. The heritability is then, according to the definition, $\frac{\sigma_g^2}{\sigma_p^2}$. *Fisher* [1951] states that the coefficient of heritability is "one of those unfortunate short-cuts, which have often emerged in biometry for lack of a more thorough analysis of the data." A warning may be needed against any uncritical calculation of this coefficient. However, where the play with gene symbols has proved to be futile under the present conditions, other methods ought to be tried, even when it is realized that these methods may yield only rough approximations which must be considered as provisional.

Generally two methods of approach are used by animal geneticists in investigations on the relative influence of heredity and environment on the variation of quantitative characters in cattle, viz. (1) experiments with monozygous twins and (2) statistical analyses of unselected data from herds and breeds. Each line of approach has its advantages and its limitations, the more important of which will be pointed out here.

(1) In an experiment with identical cattle twins it is possible to measure the phenotypic effect of a difference in one or several strictly controlled environmental factors thus demonstrating the relative modifiability of various quantitative characters of the same animals (cf. *Bonnier* and *Hansson* [1948]). By keeping two sets of identical

twins on clearly different planes of nutrition it should be possible also to study the importance of non-linear interactions between genotype and environment which is very difficult to do in a satisfactory way by other means. However, heritability in the narrow sense of the term cannot be calculated from identical-twin data for the following reasons:

(a) The variance between pairs of monozygous twins contains the components caused by dominance and gene interactions (σ_d^2 and σ_i^2), but within twin pairs there are no deviations from this source.

(b) The two members of a twin pair develop in the same uterus at the same time, and are contemporary also in their post-natal life. Therefore, the environmental variation between them will be much less, on an average, than between other animals of the same age, drawn from the same herd.

By keeping identical twins in a controlled environment it is possible to show to what extent animals of the same genotype repeat the yield of each other under the existing conditions. The repeatability thus obtained is usually higher than when it is calculated on an intra-cow basis from the data of the cow testing associations because of the contemporaneousness of yield, as well as the more efficient yield-test, and the better supervised feeding of the cows in an experimental herd than under ordinary farm conditions. Heritability figures should be calculated on the basis of a sample of the same population, fed, managed and tested under approximately the same conditions as the herds where the figures will be applied in the practical breeding work.

(2) In a statistical analysis heritability is measured by the extent to which related individuals are more alike than unrelated ones in a random-bred population. However, it is extremely difficult to separate the variance caused by the additive effect of the genes from the environmental variance, because many of the environmental factors are not distributed at random, e.g. there may be a certain trend in the plane of nutrition of a certain herd, and there are usually great environmental differences between the various herds. In practice, the matings do not take place at random, but the degree of inbreeding and the intensity of selection are generally rather low, and from our point of view the studied populations may, therefore, be considered as approximately random bred. It is obvious that the errors involved in the estimates of heritability increase with a decrease in the genetic relationship between the animals used, e.g. it

is higher for half-sib groups than for full-sibs or dam-daughter pairs. Since the coefficient of heritability is a ratio, its value can change as either the numerator or the denominator changes. Increased environmental variation may increase the denominator, thus lowering heritability of a population which has not changed genetically, whereas the breeding methods applied may change the genetic variance. In both cases the total variance may be influenced not only by changes in the strictly genetic and environmental variance but also by changes in the non-linear interactions between genes and environment. It may be expected, therefore, that different values of heritability would be obtained for the same quantitative character, e.g. milk yield or butterfat percentage, when studied under different environmental conditions, or on different populations of dairy cattle. The purpose of the present paper is to throw some light on the question how different management and plane of nutrition may influence the manifestation and heritability of quantitative characters in dairy cattle.

Material and methods.

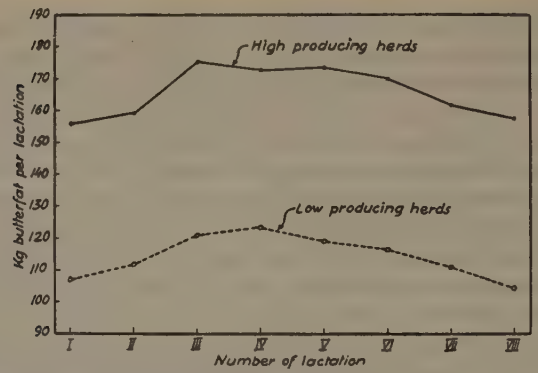
The investigations are based on the official records from all the cows in 12 low and 17 high producing herds of Swedish red and white cattle (SRB). The average yield of the cows which had completed their five first lactations in these herds (412 cows in the low and 831 cows in the high producing herds), as well as the average length of their dry periods and calving intervals, is shown in table 1. None of the herds in the low group had a higher average yearly record than 124 kgs of butterfat and no herd in the high group had a lower average than 145 kgs. The average age of the cows at the first calving was approximately the same in both groups, viz. 2 years and 11 months, and 2 years and 10 months respectively. The dry period was, on an average, somewhat longer and the calving interval somewhat shorter in the low than in the high producing herds but the differences were small and had, therefore, only a slight effect on the milk records. The herds in both groups descend from the same stock, and bulls of the same or related breeding lines have been used. The main difference between the two groups would seem to be that the bulls in the high group come closer to the "ideal type" of the breed as far as conformation is concerned. In addition, the intensity of culling on the production of young cows has probably been somewhat more pronounced in the high than in the low group.

The analysis of the data is based on the butterfat yield per lactation period, comprising not more than 300 days from calving, and on the butterfat percentage of the milk produced in the same time. The total butterfat is used as a measure of the quantity of yield; approximately the same results would be obtained by using the total amount of milk (cf. *Johansson* and *Hansson* [1940], pp. 9–10). The fat percentage is a measure of the quality of the milk as far as its fat content is concerned. The length of calving interval is used as an expression of the regularity of breeding, and the length of dry period as an indication of the persistency in yield. Two series of calculations have been made, viz. (1) an analysis of variance of the records of all cows, which have completed their first five lactations, in order to show the differences between herds and between cows within herds, and (2) an analysis of co-variance between daughters and their dams in the butterfat yield during the first lactation, and the fat percentage of the milk, with elimination of the differences between herds and between groups of daughters from the same sire.

The estimates of the theoretical components of the total variance have been made according to the method developed by *Winsor* and *Clarke* [1940].

Results of the analysis of the data.

Such characters as the milk and butterfat yield per lactation, the dry period and the calving interval may be measured several times in the same animal, i.e. after each calving. In a study of the relative importance of heredity and environment for the variation in these characters it would seem to be a logical procedure to start with a separation of the variance "within cows" from the variance between herds and "between cows within herds". The variation in yield of the same cow from one lactation to another is wholly non-genetic. Partly it is due to the change in capacity of yield with age (cf. fig.), but most of the intra-cow variance arises from controllable and non-controllable environmental influences. There are good reasons to assume that only a very small part of the variance in total milk or butterfat between herds is genetic, because the differences in feeding and management are generally rather marked, whereas the average relationship between the animals within the same herd is not much closer than that between herds of the same breed. A large part of the variance between contemporary cows within the same herd may be expected to be of genetic origin.



The relation between age and butterfat yield of cows which all had completed their 8 first lactations: averages for 308 cows in high producing and 145 cows in low producing herds.

Table 1 presents the results of such an analysis of the data from low and high producing herds. The total variance in butterfat yield is larger in the high than in the low producing group but the relative variation (the coefficients of variation) is smaller. The variance between herds is most pronounced for the butterfat percentage, probably due to differences in the intensity of selection for this character, and it is very small for the length of the calving interval. The effect of age on yield is somewhat more pronounced in the high than in the low group; 22.0 and 14.6 per cent respectively of the variance "within cows" may be attributed to this cause. There is no effect of age on butterfat percentage and the effect on calving intervals is very slight. If the variance between cows within herds is denoted σ_b^2, the variance within cows σ_w^2, and the variance owing to changes in age from the first to the fifth lactation σ_a^2, the *repeatability* of the characters is best expressed by the fraction $\frac{\sigma_b^2}{\sigma_b^2 + (\sigma_w^2 - \sigma_a^2)}$. The coefficients of repeatability in the four characters are only slightly lower for the herds on low yield than for the herds on high yield. In this case, therefore, the individuality (or genotype) of the animal has manifested itself just as clearly in a "poor" as in a "good" environment, although there was a pronounced difference in the level of butterfat yield. The repeatability is high for butterfat percentage (0.594 and 0.640 respectively), lower for total yield (0.412 and 0.432),

Table 1. Distribution of the variance in butterfat yield, fat content of the milk, length of dry period and length of calving interval for two genetically similar groups of cows under different environmental conditions.

Sources of variance	Herds with low yield					Herds with high yield				
	Degrees of freedom	Mean square, estimated components of variance, etc.				Degrees of freedom	Mean square, estimated components of variance, etc.			
		Butterfat yield	Butterfat percentage	Dry period	Calving interval		Butterfat yield	Butterfat percentage	Dry period	Calving interval
Mean square	2059	840.3	0.086	2037.4	4727.0	4154	1296.9	0.097	1595.9	3553.6
Relative variance: %										
Between herds . . .	10	4.9[3]	21.6[3]	3.9[3]	0.9[2]	16	7.1[3]	23.5[3]	7.7[3]	1.2[3]
Between cows within herds	401	35.6[3]	46.6[3]	19.1[3]	3.3[3]	814	34.6[3]	49.0[3]	17.9[3]	6.2[3]
Within cows	1648	59.5	31.8	77.0	95.8	3324	58.3	27.6	74.4	92.6
Variance within cows:										
Between lactations, %	4	14.6[3]	—	4.5[3]	2.4[3]	4	22.0[3]	—	15.9[3]	1.5[3]
Residual, %	1648	85.4	100.0	95.5	97.6	3320	78.0	100.0	84.1	98.5
Repeatability within herds (age corrected)[1]		0.412[3]	0.594[3]	0.206[3]	0.035[3]		0.432[3]	0.640[3]	0.223[3]	0.063[3]
Standard deviation (σ)										
Total		29.0	0.293	45.1	68.8		36.0	0.311	40.0	59.6
Within cows (age corrected)[1]		20.7	0.167	38.8	66.5		24.4	0.164	31.7	56.9
Coefficient of variation		25.4	7.34	57.9	17.9		22.2	7.77	59.9	15.5
Average of the characters studied	Number of cows 412	kgs 114.0	% 3.99	days 79.0	days 382.9	Number of cows 831	kgs 160.0	% 4.00	days 66.8	days 384.0

[1] The variance "between lactations", calculated from the averages of the first, second, third, fourth and fifth lactations, is subtracted from the variance within cows.
[2] = $P < 0.01$. [3] = $P < 0.001$

Table 2. Dam (x)-daughter (y) correlations (r) and regressions (b) in regard to butterfat yield and butterfat percentage: first lactation record (300 days).

Sources of covariance	Herds with high yield					Herds with low yield				
	D.F.	Butterfat yield		Butterfat percentage		D.F.	Butterfat yield		Butterfat percentage	
		r_{xy}	$b_{y/x}$	r_{xy}	$b_{y/x}$		r_{xy}	$b_{y/x}$	r_{xy}	$b_{y/x}$
Total	1050	0.230[3]	0.234	0.426[3])	0.437	2397	0.253[3]	0.275	0.415[3]	0.395
Between herds . . .	11	0.851[3]	0.524	0.902[3]	0.862	18	0.604[2]	0.618	0.730[3]	0.565
Within herds	1038	0.127[3]	0.140	0.254[3]	0.267	2378	0.195[3]	0.215	0.325[3]	0.329
Between sires	34	−0.068[0]	−0.137	0.145[0]	0.219	63	0.297[1]	0.370	0.133[0]	0.185
Within daughter-dam groups (= within sires)	1003	0.160[3]	0.162	0.268[3]	0.271	2314	0.182[3]	0.196	0.350	0.342

[0] = $P > 0.05$ [1] = $P < 0.05$ [2] = $P < 0.01$ [3] = $P < 0.001$

still lower for length of the dry period (0.206 and 0.223) and very low for length of the calving interval (0.035 and 0.063). The irregularity in pregnancies depends, apparently, almost wholly on non-genetic factors. It should be kept in mind, however, that all cows in this analysis had completed at least five lactations. Poor breeders culled at earlier stages were not included.

Table 2 shows the result of the analysis of co-variance between dams and daughters in regard to butterfat yield and butterfat percentage during the first lactation, with elimination of the differences between herds and between groups of fraternal half sisters (= between sires) within the same herd. The regression of the daughters of the same sire on their dams depends almost exclusively on the additive effect of the genes. By eliminating the co-variance between herds and between groups of fraternal half sisters, the dam-daughter correlation will be fairly well isolated from its environmental component. The heritability (h^2) of the character may, therefore, be estimated by doubling the intra-sire regression of daughters on dams ($h^2 = 2\, b_{y/x}$). By this doubling the following coefficients of heritability are obtained:

	Herds with low yield	Herds with high yield
Total butterfat	0.32	0.39
Butterfat percentage	0.54	0.68

The figures are a little higher for herds on a high level of yield than for the low producing herds but the differences were found not to be significant. The P-values of the difference between the regression coefficients is for butterfat yield $0.4 > P > 0.3$ and for butterfat percentage $0.1 > P > 0.05$.

The figures obtained here for repeatability and heritability agree very well with those found by investigators in Great Britain and U.S.A. (cf. review by *Johansson* [1950]). *Mahadevan* [1951] obtained the following results in an analysis of data for the Ayrshire cattle in Scotland.

	Repeatability within herds, after correction for age, season of calving and length of calving interval	Heritability
Milk yield	0.47	0.31
Butterfat percentage	0.70	0.56

It is rather surprising that the figures obtained for different breeds living under different environmental conditions, agree so closely with one another.

Reference may be made to *Falconer* and *Latyszewski* [1952] who selected mice for body weight at 6 weeks of age on two planes of nutrition and found that the heritability was higher in the restricted diet strain (though not significantly so) than in the strain fed ad libitum. It is pointed out that the results do not support *Hammond*'s [1947] thesis that "the character required is best selected for under conditions which favour its fullest expression", an opinion which is generally held also by the practical breeders. *Falconer* and *Latyszewski*'s results are of great interest in this connection, and they agree on the whole with our own results.

The theoretical and practical importance of the variation caused by interactions between heredity and environment would justify large scale experiments for further elucidation of this problem. In regard to the production characters of dairy cattle monozygous twins are best suited for such experiments.

Summary.

An analysis has been made of the milk records from two groups of herds of Swedish red and white cattle, one group with a low and one with a high average yield. It is considered that the genetic differences between the two groups were very small, on an average, and that the difference in yield was due, almost ex-

clusively, to differences in feeding and management. The object was to study the repeatability of yield of the same animals from one lactation to another, as well as the heritability within genetically similar groups on two different planes of environment. The results are presented in tables 1 and 2.

The repeatability and heritability of the butterfat yield per lactation, and the butterfat percentage, were only slightly lower for the low producing than for the high producing herds. The results, therefore, do not support the general opinion that genetic differences in regard to quantitative characters in our domestic animals are more clearly manifested in an optimum environment than under poorer conditions of feeding and management.

Résumé.

L'auteur a analysé le rendement de lait dans deux troupeaux de bétail suédois de la race rouge et blanche, dont un groupe avait un rendement moyen bas et l'autre un rendement moyen élevé. On peut estimer que les différences génétiques entre les deux groupes étaient peu considérables en moyenne et que les différences de rendement dépendaient presque exclusivement de l'alimentation et du traitement différents. On avait l'intention d'étudier la répétition du rendement dans les mêmes animaux pendant une période s'étendant d'une lactation jusqu'à la suivante ainsi que la transmissibilité dans des groupes similaires d'un point de vue génétique et placés dans deux milieux différents. Les résultats obtenus ressortent des tableaux 1 et 2.

La répétition et la transmissibilité du rendement de graisse par lactation, ainsi que le pourcentage de graisse n'étaient que légèrement plus bas dans les troupeaux ayant un rendement bas que dans ceux qui avaient un rendement élevé. Donc, les résultats ne viennent pas à l'appui de l'opinion générale selon laquelle les différences génétiques de caractères quantitatifs dans nos animaux domestiques se manifestent plus nettement dans les conditions de milieu les plus favorables que dans des conditions d'alimentation et de traitement moins bonnes.

Zusammenfassung.

Der Verfasser hat die Milchproduktion bei zwei Gruppen Bestände des schwedischen Rot-Weiß-Viehes analysiert, von welchen die eine Gruppe niedrige, die andere hohe durchschnittliche Milchproduktion hatte. Man dürfte annehmen, daß die genetischen Unterschiede zwischen den beiden Gruppen einen geringen Durchschnittswert haben und daß die Unterschiede in der Milchproduktion fast ausschließlich auf Unterschiede in Fütterung und Pflege zurückzuführen sind. Der Zweck der Untersuchung war, die Wiederholung der Milchleistung desselben Tieres von einer Laktation zur anderen zu studieren sowie die Hereditabilität innerhalb genetisch gleichgearteter Gruppen unter zwei verschiedenen Milieuverhältnissen. Die Ergebnisse gehen aus den Tabellen 1 und 2 hervor.

Wiederholung und Hereditabilität von Butterfettproduktion per Laktation, sowie der Prozentsatz an Butterfett, waren nur um weniges geringer für die niedrigerzeugenden, verglichen mit den hocherzeugenden Beständen. Die Ergebnisse unterstützen daher keineswegs die allgemeine Auffassung, daß sich genetische Unterschiede im Hinblick auf qualitative Eigenschaften bei unseren Haustieren deutlicher unter optimalen Milieuverhältnissen, verglichen mit schlechteren Fütterungs- und Pflegebedingungen, manifestieren.

LITERATUR

Bonnier, G. and *A. Hansson:* Hered. *2*, 1–27, 1948. – *Falconer, D. S.* and *M. Latyszewski:* J. Genet. *51*, 67–80, 1952. – *Fisher, R. A.:* Brit. agric. Bull. *4*, 217–218, 1951. – *Hammond, J.:* Biol. Rev. *22*, 195–213, 1947. – *Johansson, I.:* Anim. Breed. Abstr. *18*, 1–12, 1950. – *Johansson, I.* and *A. Hansson:* Causes of Variation in Milk and Butterfat Yield of Dairy Cows. K. Lantbr. Akad. Handl. Stockholm 1940. – *Lush, J. L.:* Heritability of Quantitative Characters in Farm Animals. Proc. 8th inter. Congr. Genet. Hered. Suppl. 356–375, 1949. – *Mahadevan, P.:* J. agric. Sci. *41*, 80–88, 94–97, 1951. – *Winsor, C. P.* and *G. L. Clarke:* J. Marine Research *3*, 1–34, 1940.

Josephson, B.: Acta genet. *4*, 231–235, 1953

From the Central Clinical Laboratory, St. Erik's Hospital, Stockholm

THE ICTERUS INDEX AS A MEASURE OF THE SERUM BILIRUBIN CONCENTRATION

By BERTIL JOSEPHSON

Since *Meulengracht* [1920] described his method for comparing the intensity of the colour of plasma or serum with that of a standard dichromate solution the method has been extensively used in routine clinical work as an estimation of the degree of icterus. Previously *Nyström* [1915] had described a similar method for estimating the serum colour by comparison with a solution of $FeCl_3$ containing a trace of safranin. *Nyström* was aware of the influence of the bile pigments on the colour but he was interested only in the "normal" colour in patients without jaundice.

The presence of bile pigments in the blood was first demonstrated in horse serum by *O. Hammarsten* [1878]. The corresponding demonstration in normal human serum was published by *Hannema* [1915]. It is well known that beside the bile pigments several coloured substances are present in the normal human plasma such as lutein, "lipochrom" carotenes, etc. Their concentration, however, is —except under rare, pathological conditions—so low that their

influence on the colour of plasma from jaundiced patients is considered to be negligible. On the other hand, the influence of these substances on the plasma colour of non-icteric subjects in comparison to the influence of the bile pigments is not known. Only one statistical comparison between icterus index and chemically determined bilirubin seems to have been published (*Sterner* and *Cusack* [1945]). Their material was 429 patients with icterus indices from normal up to 25. The results from icteric and non-icteric subjects were treated in common. They determined the icterus index after precipitation of the serum proteins with acetone which is known to co-precipitate a very considerable part of the bile pigments (*Newburger* [1937]).

As the determination of the icterus index is much less complicated and takes a shorter time to carry out than the determination of the bile pigments by the diazo reaction it was considered useful to examine the influence of the concentration of these pigments on the icterus index in the healthy man. This influence was compared with that in jaundiced patients.

Material and technique.

196 healthy men, 164 healthy women and 40 jaundiced patients[1] were examined (cf. table 1). 193 of the healthy subjects were less than 50 years old and 167 were over 60 years. The normal subjects and the laboratory results from these were the same as those used by *Josephson* and *Dahlberg* [1952] and the control of the health of the subjects is described in their paper. Icterus index was determined according to *Meulengracht* [1921] and bilirubin according to *Jendrassik* and *Grof* [1938]. Both determinations were carried out on serum. In neither method are the proteins precipitated. The diagnoses, age and sex of the icterus patients were not considered to be of interest.

Statistical methods.

The means, their standard errors and the standard variations are calculated as in the paper of *Josephson* and *Dahlberg* [1952]. The correlation coefficients and regression lines are calculated according to *Bravais-Pearson* as described by *Dahlberg* [1940][2].

[1] Dr. *J. Ström*, physician in chief of the Hospital for Epidemic Diseases of Stockholm, has kindly provided me with serum samples from patients with epidemic hepatitis. Some sera were from patients with non-infectious liver diseases treated in St. Erik's Hospital.

[2] I owe many thanks to Dr. *E. Lander* who has helped me with the statistical treatment of the results in a very kind way.

Table 1. The mean values, M, the standard errors of the means, ε(M), and the standard variations, σ, of the icterus index and the bilirubin concentration. The coefficients, r, and their standard errors, ε(r), for the correlation between icterus index and bilirubin concentration.

State of health	Sex	Age years	Number	Icterus index		Bilirubin		Correlation coefficient r ± ε(r)
				M ± ε(M)	σ	M ± ε(M)	σ	
healthy	men	< 50	83	6,45±0,25	2,27	0,929±0,053	0,483	0,83±0,034
		> 60	113	5,46±0,16	1,66	0,740±0,029	0,311	0,64±0,056
		total	196	5,88±0,14	2,00	0,820±0,029	0,404	0,77±0,029
healthy	women	< 50	110	5,80±0,12	1,29	0,788±0,026	0,276	0,53±0,069
		> 60	54	4,37±0,13	0,97	0,766±0,033	0,244	0,59±0,089
		total	164	5,33±0,11	1,37	0,780±0,021	0,266	0,49±0,059
healthy	both	< 50	193	6,08±0,13	1,81	0,848±0,028	0,386	0,75±0,032
		> 60	167	5,10±0,12	1,56	0,749±0,023	0,292	0,58±0,052
		total	360	5,63±0,09	1,76	0,802±0,018	0,349	0,69±0,027
jaundiced	both		40	35,1±6,4	40,6	5,36±1,10	6,98	0,84±0,047

Results.

The means, variations and correlation coefficients are demonstrated in table 1. The results from the healthy men are plotted in fig. 1 and from the jaundiced patients in fig. 2. The results from the healthy women

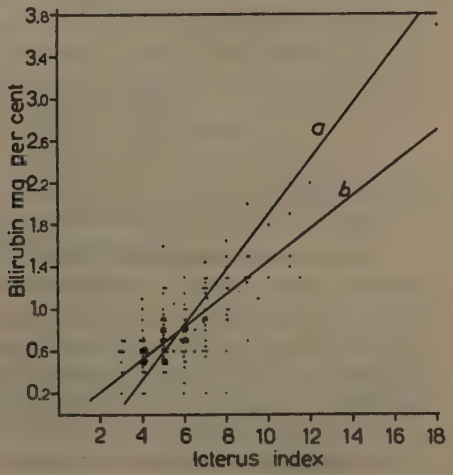


Fig. 1. The icterus indices and bilirubin concentrations of the 196 healthy men. The line *a* represents the icterus index as a function of the bilirubin concentrations and the line *b* the bilirubin concentration as a function of the icterus indices.

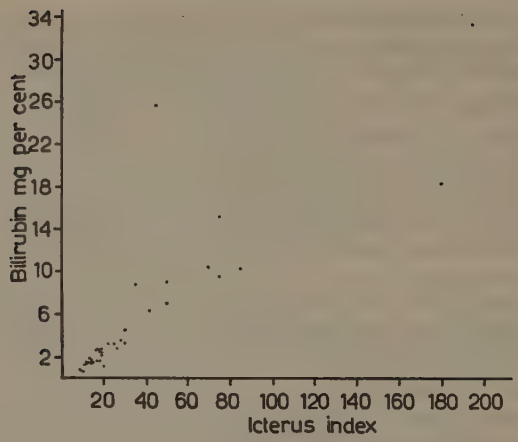


Fig. 2. The icterus indices and bilirubin concentrations of 40 jaundiced patients.

are not recorded as they did not differ very much from those of the men. In fig. 1 the regression lines are drawn showing the icterus index as a function of the bilirubin concentrations (a) and the bilirubin concentration as a function of the icterus indices (b).

The calculation of these lines gave the following results for the mutual dependence of icterus index (I) and bilirubin concentration (B) on each other.

For men $I = 3{,}81\,B + 2{,}76 \pm 1{,}28$
$B = 0{,}155\,I - 0{,}09 \pm 0{,}26$

For women $I = 2{,}53\,B + 3{,}35 \pm 1{,}20$
$B = 0{,}095\,I + 0{,}27 \pm 0{,}23$

Discussion.

As the table demonstrates, there was a fairly good correlation between icterus index and serum bilirubin concentration in all the groups of normal subjects. The correlation coefficient was markedly high in the group of younger men. It was rather good in the icteric patients too. This was to be expected, as the discoloration of their serum must have been due to its high concentration of bile pigments.

The fairly good correlation in the healthy subjects shows that the bile pigments have considerable influence on the serum colour even

under normal conditions. On the other hand the correlation is not so good that one is permitted to draw any definite conclusions on the bilirubin concentration of a serum sample from its index figure.

A better correlation was to be expected in the jaundiced patients than the one obtained. The fact that the correlation coefficient in this group was only 0,84 may be explained partly by the obvious possibility of individual errors in estimation of the icterus index and partly by the fact that several different bile pigments, not identical with bilirubin, may appear in the serum of jaundiced patients. A few of the icteric sera were turbid which is the probable reason why their figures went very much astray.

The fact that the regression lines for the icterus indices as a variable of the bilirubin concentrations do not point towards the origin but extrapolate towards a value of 2 to 4 at a bilirubin concentration of *Zero* is of interest. This observation indicates that approximately 2 to 4 units of the index figures are due to coloured substances in the serum, which are not bilirubin.

Summary.

The correlation between the icterus index according to *Meulengracht* [1920] and the serum bilirubin concentration was examined in 360 healthy subjects and in 40 jaundiced patients. In the normal subjects the correlation coefficient was 0,69 and in the patients it was 0,84.

Résumé.

La corrélation entre l'index d'ictère selon *Meulengracht* [1920] et la concentration de bilirubine dans le sérum a été analysée chez 360 personnes bien portantes et chez 40 malades ayant la jaunisse. Chez les personnes normales le coefficient de corrélation était de 0,69 et chez les malades de 0,84.

Zusammenfassung.

Die Korrelation zwischen dem Ikterusindex nach *Meulengracht* [1920] und der Konzentration von Bilirubinserum ist bei 360 gesunden Individuen und 40 Gelbsuchtpatienten untersucht worden. Für die normalen Individuen betrug der Korrelationskoeffizient 0.69 und für die Patienten 0.84.

REFERENCES

Dahlberg, G.: Statistical Methods, London 1940. – *Hammarsten, O.:* Malys Jb. Thier Chem. *8*, 129, 1878. – *Hannema, L. G.:* Malys Jb. Thier Chem. *45*, 88, 1915. – *Jendrassik, L.* und *P. Grof:* Biochem. Z. *297*, 81, 1938. – *Josephson, B.* and *G. Dahlberg:* Scand. J. clin. Lab. Invest. *4*, 216, 1952. – *Meulengracht, E.:* Dtsch. Arch. klin. Med. *132*, 285, 1920; Ugeskr. Laeg. *20*, 1, 1921. – *Newburger, R.:* J. Lab. clin. Med. *22*, 1192, 1937. – *Nyström, B.:* Finska LäkSällsk. Månadst. *57*, 1649, 1915. – *Sterner, J.* and *M. Cusack:* Amer. clin. Path. *15* (techn. sect. 9), 7, 1945.

Kallós, P. und *L. Kallós-Deffner*: Acta genet. *4*, 236–240, 1953

Aus dem Margareta und Erik Wihlborg Laboratorium für Allergieforschung
Hälsingborg, Schweden

EXPERIMENTELLE BEITRÄGE ZUM PROBLEM DER VERERBUNG DER ALLERGISCHEN DISPOSITION

Von PAUL KALLÓS und LISELOTTE KALLÓS-DEFFNER

G. Dahlberg hat 1939 die Literatur über die Vererbung der allergischen Disposition einer Analyse unterzogen. Er ist dabei zu der Schlußfolgerung gekommen, daß man wohl berechtigt ist anzunehmen, daß «eine erbliche Anlage eine gewisse Rolle» für das Auftreten von allergischen Krankheiten spielt, daß man jedoch keine klare Vorstellung darüber erlangen kann «wie groß die Rolle ist, die die Erbanlage spielt». Diese Schlußfolgerung besteht immer noch zu Recht. Aus der Darstellung von *G. Dahlberg* geht auch klar hervor, wie groß die Schwierigkeiten sind, falls man die Bedeutung der erblichen Anlage bei allergischen Patienten ermitteln will. Bei der experimentellen Allergisierung von Versuchstieren liegen natürlich wesentlich einfachere Verhältnisse vor, es ist jedoch bisher nicht gelungen, den Nachweis zu erbringen, daß die Unterschiede der Allergisierbarkeit, welche zwischen den Individuen ein- und derselben Spezies zu beobachten sind, durch Erbfaktoren bedingt werden (*Bret Ratner* und *D. E. Silberman* [1953]). Vergleicht man verschiedene Tierarten im Hinblick auf die Allergisierbarkeit miteinander, so ergibt sich, daß unter denselben Versuchsbedingungen Meerschweinchen sehr leicht, Ratten und Mäuse hingegen sehr schwer zu allergisieren sind. Eine einzige Injektion einer minimalen Dosis eines artfremden Eiweißkörpers führt bei praktisch genommen allen Meerschweinchen dazu, daß die intravenöse Injektion einer ebenfalls sehr kleinen Menge desselben Eiweißkörpers wenigstens eine Woche nach der Erstinjektion verabreicht, einen tödlichen anaphylaktischen Schock auslöst. Bei Mäusen und Ratten benötigt man zur Allergisierung wiederholte Injektionen von relativ großen Allergendosen und bei der Auslösung treten meistens nur leichtere anaphylaktische Symptome auf oder gehen die Tiere im protrahierten Schock ein

(*R. S. Weiser* und Mitarbeiter [1941], *Ph. D. McMaster* [1941], *Ph. D. McMaster* und *St. Hudack* [1932], *Ph. D. McMaster* und *H. Kruse* [1949], *R. L. Mayer* und *D. Brousseau* [1946]). Dieser Unterschied zwischen den genannten Tierarten, welcher bei allen Individuen zutage tritt, wird allgemein als erblich bedingt aufgefaßt (*D. Perla* und *J. Marmorston* [1941]). Interessanterweise ergibt sich eine gewisse Parallelität zwischen der Allergisierbarkeit und der Empfindlichkeit für Histamin. Die tödliche Histamindosis (intravenöse Verabreichung) beträgt für Meerschweinchen 0,3 mg/kg, für Mäuse und Ratten 500–750 mg/kg. Diese natürliche Resistenz der Mäuse und Ratten gegen Histamin wird ebenfalls als erblich bedingt angesehen. *I. A. Parfentjev* und Mitarbeiter haben 1947 die Beobachtung gemacht, daß Mäuse, welche eine intraperitoneale Injektion von H. pertussis-Vaccine erhalten haben, 4 Tage hernach etwa 25–50mal empfindlicher für Histamin sind als unvorbehandelte Tiere. Die Erhöhung der Empfindlichkeit für Histamin konnte nur durch H. pertussis-Vaccine oder durch das Überstehen einer Infektion mit H. pertussis (*M. Pittman* [1951]) herbeigeführt werden, alle anderen Bakterienarten, welche bisher geprüft worden sind, haben sich als unwirksam erwiesen. *I. A. Parfentjev* hat gleichzeitig auch den Befund erhoben, daß die vorbehandelten Mäuse nicht nur für Histamin empfindlich werden, sondern daß man bei ihnen den anaphylaktischen Schock mit geringen Mengen von H. pertussis-Vaccine bzw. -Extrakt auslösen kann, was den Eindruck erweckt, daß die Tiere durch H. pertussis-Vaccine hochgradig allergisch gemacht werden. *I. A. Parfentjev* und Mitarbeiter haben nicht untersucht, ob mit H. pertussis vorbehandelte Mäuse auch mit anderen Allergenen leichter allergisiert werden können. Jedenfalls handelt es sich hier darum, daß eine für eine Tierart charakteristische erblich bedingte Eigenschaft durch einen Milieufaktor, nämlich durch eine Infektion oder Vorbehandlung mit abgetöteten Bakterien einer bestimmten Art, entscheidend verändert werden kann. Ähnliche Einflüsse können sich vielleicht auch bei der Allergisierung anderer Tierarten geltend machen. Dies hat uns veranlaßt, die Angaben von *I. A. Parfentjev* und Mitarbeiter, die übrigens bezüglich der Histaminempfindlichkeit von *B. N. Halpern* und *J.-L. Roux* [1950] bestätigt worden sind, nachzuprüfen. Wir haben uns jedoch in Anlehnung an unsere früheren Arbeiten (*P. Kallós* und *W. Pagel* [1937], *P. Kallós* und *L. Kallós-Deffner* [1939]) einer anderen Versuchsanordnung bedient und bei der Auslösung der Histaminvergiftung bzw. anaphylaktischen Erschei-

nungen den wirksamen Stoff als Aerosol zugeführt. Hier soll über unsere Ergebnisse in aller Kürze berichtet werden.

Alle Versuche wurden an weiblichen weißen Mäusen im Gewicht von 20 g durchgeführt. Alle Mäuse gehörten dem Inzuchtstamm der Isko-Farm in Pataholm an. Bei unvorbehandelten Mäusen dieses Stammes löst eine halbstündige Einatmung von Histaminaerosol (2–5 % Histamindiphosphatlösung) oder Eiereiweißaerosol (10 %) keinerlei Erscheinungen aus, auch dann nicht, falls die Einatmung in kürzeren oder längeren Zeitabständen wiederholt wird. Sensibilisiert man die Mäuse durch eine einmalige intraperitoneale Injektion von 1–10 mg Eiereiweiß und exponiert diese Tiere 4–7–14 Tage später für Histamin- bzw. Eiereiweißaerosol, so verhalten sie sich wie unvorbehandelte Tiere.

Die Vorbehandlung von Mäusen durch eine einmalige Injektion (intraperitoneal) von 10 000 Millionen Phase I Pertussisbakterien (Vaccine von Statens Bakteriologiska Laboratorium, Stockholm) verändert ihre Histaminempfindlichkeit in hohem Grade. Bei Exposition für Histaminaerosol 4 Tage nach der Vorbehandlung treten nach 5–15 Minuten langer Einatmung schwere Krankheitserscheinungen auf, die bei etwa der Hälfte der Tiere tödlich enden. Der Tod erfolgt in tonisch-klonischen Krämpfen, denen Unruhe, eine eigenartige Haltung (die Tiere sitzen auf den Hinterbeinen, gestützt auf die Vorderpfoten, wobei der Kopf maximal nach aufwärts gestreckt wird) Atemnot und hochgradige Cyanose vorausgingen. Während des ganzen Verlaufes scheint intensiver Juckreiz zu bestehen, welcher die Tiere zwingt, sich intensiv zu kratzen.

Wird die Vorbehandlung so vorgenommen, daß die Mäuse außer Pertussis-Vaccine auch Eiereiweiß intraperitoneal erhalten, und zwar in der oben genannten Dosierung, kann man bei ihnen nach 1–2 Wochen die oben geschilderten Erscheinungen mit Eiereiweißaerosol auslösen. Die Erscheinungen sind mit denen, welche durch Histaminaerosol ausgelöst werden können, identisch, wobei jedoch Unterschiede in der Intensität bestehen. Es werden nicht alle Mäuse gleich stark anaphylaktisch. Der Juckreiz ist auch in dieser Versuchsreihe konstant zu beobachten gewesen und es soll hier erwähnt werden, daß diese Versuchsanordnung u. E. geeignet ist, um die Wirksamkeit von juckstillenden Stoffen zu messen.

Die Vorbehandlung mit H. pertussis-Vaccine macht somit die Mäuse empfindlicher für Histamin und leichter allergisierbar. Die «allergische Disposition» wird nicht nur für die Antigene des H.

pertussis erhöht, sondern auch für gleichzeitig zugeführte Eiweißkörper anderer Herkunft. Bekanntlich wird angenommen, daß wenigstens ein Teil der allergischen Erscheinungen bei Meerschweinchen, Menschen und anderen Tierarten darauf zurückzuführen ist, daß bei der zellständigen Reaktion von Allergen und Antikörper Histamin freigemacht wird (*H. Dale* [1948]). Nach *I. A. Parfentjev* und Mitarbeiter beträgt die intravenöse tödliche Histamindosis für durch H. pertussis maximal empfindlich gemachte Mäuse immer noch etwa 20 mg/kg. Der Histamingehalt des Mäuseorganismus wird mit 10 mg/kg angegeben, weswegen es sehr fraglich erscheint, ob das Freiwerden von Histamin auch bei der Maus an der Entstehung von allergischen Krankheitserscheinungen ursächlich beteiligt ist. Diese Zusammenhänge müssen noch näher erforscht werden.

Die Ergebnisse, welche hier kurz angeführt worden sind, sprechen dafür, daß die erblich bedingte allergische Disposition keine unveränderliche Größe darstellt, sondern durch Milieufaktoren beeinflußt werden kann. Weitere Untersuchungen in dieser Richtung werden vielleicht dazu beitragen können die Rolle der Erbfaktoren im Sinne von *G. Dahlberg* näher aufklären zu können.

Zusammenfassung.

Die Resistenz von weißen Mäusen gegen Histaminvergiftung und Allergisierung, die allgemein als erblich bedingt angesen wird, kann, in Übereinstimmung mit den Angaben von *I. A. Parfentjev*, durch eine einmalige intraperitoneale Zuführung von H. pertussis Phase I-Vaccine erheblich vermindert werden. Die mit Vaccine behandelten Mäuse erkranken bei Einatmung von Histaminaerosol mit charakteristischen Symptomen. Sie können mit Eiweißallergenen regelmäßig allergisch gemacht werden und dieselben Symptome, welche die Histaminvergiftung charakterisieren, können bei allergischen Mäusen durch Einatmen des spezifischen Allergenaerosols ausgelöst werden.

Résumé.

La résistance des souris blanches à l'intoxication par l'histamine et à l'action d'allergiser, laquelle est généralement considérée comme héréditaire, peut être considérabelment réduite, selon les rapports de *I. A. Parfentjev*, par une seule injection intra-péritonéale de H. pertussis Phase I-Vaccine. Les souris inoculées tombent malades en montrant des symptômes caractéristiques quand elles respirent de l'aérosol de l'histamine. Elles peuvent en règle générale devenir allergiques sous l'effet d'allergènes d'albumen et les mêmes symptômes que ceux qui caractérisent l'intoxication par l'histamine peuvent se produire si elles respirent de l'aérosol de l'allergène spécifique.

Summary.

The resistance to histamine poisoning and experimental allergic reactions which is generally believed to be genetically determined could be considerably reduced by one single intraperitoneal injection of H. pertussis Phase I-Vaccine. This is in accordance with the reports of *I. A. Parfentjev*. Mice treated with this vaccine display characteristic symptoms of disease after inhalation of a histamine aerosol. They can also be made allergic by treatment with protein allergens and after inhalation of specific allergen aerosols the same symptoms as in histamine poisoning appear.

LITERATUR

Dahlberg, G.: Die Vererbung der allergischen Disposition. Fortsch. Allerg. *1*, 330, S. Karger, Basel/New York 1939. – *Dale, H.:* Brit. med. J. *1948/II*, 281. – *Halpern, B. N.* et *J.-L. Roux:* C. R. Soc. Biol., Paris *143*, 923, 1950. – *Kallós, P.* und *W. Pagel:* Acta med. scand. *91*, 292, 1937. – *Kallós, P.* und *L. Kallós-Deffner:* Die experimentellen Grundlagen der Allergielehre. Fortschr. Allerg. *1*, 5, S. Karger, Basel/New York 1939. – *Mayer, R. L.* and *D. Brousseau:* Proc. Soc. exp. Biol., N.Y. *63*, 187, 1946. – *McMaster, Ph. D.:* J. exp. Med. *74*, 29, 1941. – *McMaster, Ph. D.* and *St. Hudack:* J. exp. Med. *55*, 417, 1932. – *McMaster, Ph. D.* and *H. Kruse:* J. exp. Med. *89*, 583, 1949. – *Parfentjev, I. A.:* Yale J. Biol. Med. *23*, 28, 1950/51. – *Parfentjev, I. A.* and *M. A. Goodline:* J. Pharmacol. *92*, 411, 1948. – *Parfentjev, I. A., M. A. Goodline* and *M. E. Virion:* J. Bact. *53*, 597, 603, 613, 1947. – *Perla, D.* and *J. Marmorston:* Natural Resistance and Clinical Medicine. Little, Brown and Co., Boston 1941. – *Ratner, Bret* and *D. E. Silberman:* Trans. N. Y. Acad. Sci. Ser. II *15*, 82, 1953. – *Weiser, R. S., O. J. Golub* and *D. M. Hamre:* J. inf. Dis. *68*, 97, 1941.

Kemp, T.: Acta genet. *4*, 240–247, 1953

From the University Institute for Human Genetics, Copenhagen

GENETIC HYGIENE AND GENETIC COUNSELING

By TAGE KEMP

By Genetic Hygiene we understand the branch of medical science which aims at preventing the transmission of pathogenetic genes from one generation to the next and thereby limits their spread in the population.

Genetic Hygiene belongs to the purely medical sciences. Its sole task is that of preventing disease and consequent suffering and misery. It constitutes an essential part of preventive medicine.

Genetic hygiene corresponds in the main to negative qualitative eugenics. It rests entirely on the principles of voluntariness. Genetic-hygienic measures are taken exclusively at the wish of the persons concerned. Compulsory measures are never employed within genetic hygiene. Experience shows that patients, who have been informed of the significance of the hereditary taint, nearly always follow their doctor's advice within this field.

The patients and their relatives nearly always realize the expediency of genetic-hygienic measures and want to cooperate. Still, it is obvious that measures which interfere so radically with the fate and most intimate life of the human individual may arouse some friction or conflict of views. The physicians and other authorities dealing with eugenic cases ought always to be most considerate and thorough in their investigations, and the principle that ought always to be followed is that too few genetic-hygienic operations are preferable to too many.

In the countries where genetic-hygienic laws exist containing directions for the employment of eugenic measures, the law guides patients and doctors, and they contain securities against the misuse of the often serious measures which have to be taken.

The most important negative eugenic measures are sterilisation, induced abortion, prohibition of marriage, matrimonial advice, combined with instruction on contraception, and other forms of genetic advice, and finally segregation of the individual in a hospital or other institution.

In some countries special clinics have been established for the purpose of giving matrimonial advice. Information on contraception is usually given here as well. In Denmark matrimonial advice is frequently given by general practitioners as well as by specialists. The possibilities within this field will gradually increase with increasing medical research in heredity. Genetic counseling will in future become a comprehensive and important medical task.

It is also a matter of genetic-hygienic importance that patients with hereditary diseases should stay for a considerable length of time in hospitals or institutions, or be kept under public care. This results in a shorter or longer segregation of such individuals, which naturally lowers their chances of propagation. Thus, hospital or

institutional treatment, public care, and general knowledge of mental defectives, epileptics, the deaf, the blind, and the crippled are of genetic-hygienic importance.

The break-up of isolates where inbreeding causes recessive diseases to prevail, may be of genetic-hygienic significance. If such inbred groups are scattered throughout the population, as happens relatively often nowadays owing to improved communications, their respective hereditary diseases will cease to appear. On the other hand, the taint will become more widespread, and this may in the course of time involve a certain risk.

It is sometimes stated that the main purpose of genetic hygiene is to spare the community a good deal of expense. This is, of course, a great mistake. Negative eugenic measures are of an entirely medical character aiming at the prevention of disease and other calamities. Like so many other preventive measures in medicine they naturally mean a saving to the public in the long run.

The great development which medical genetics has undergone and the extensive knowledge of its results among physicians and the general population forms the basis for a correct execution of genetic-hygienic measures. Other requirements are, however, necessary before an efficient genetic-hygienic programme can be carried out. The public health system and social care must be well organized. Where hospitals and other institutions as well as the public care system have attained a certain standard, many patients with hereditary diseases can be traced through these institutions. Efficient social laws however, presuppose genetic-hygienic measures.

An important aid for the carrying through of genetic-hygienic measures is the establishment of a genetic-hygienic registration comprising as far as possible a complete register or card-index, continually brought up to date, of all the patients with severe hereditary diseases within a certain area, and also of their families.

A registration of this kind can be useful in many ways in genetic-hygienic questions. If in a given case the question of genetic-hygienic measures is raised, the card-index will furnish information about the person or persons in question and their families. It will thereby in many cases be easier to give advice.

In Denmark a genetic-hygienic register, a medico-genetic registration, was established in 1938 and is kept at the University Institute for Human Genetics in Copenhagen. This register is contin-

ually brought up to date covering so far as possible all the patients with severe hereditary diseases within the country.

A physician who is to make a genetic prognosis must, of course, procure information on the hereditary taint in the family. Not all physicians have realized this fact. For a proper genetic prognosis it is, however, absolutely necessary that the physician takes the, sometimes very great, trouble of procuring this information. The task of the physician in connection with genetic-hygienic measures is in the first instance to make a genetic prognosis for the offspring of a person or a married couple, in cases where information is available on diseases in one or both parents and in their families. It is not always possible to give a definite percentage figure for the chances that a child in whose family a certain hereditary disease occurs will inherit the disease.

The physician is frequently asked about the advisability of having children by patients with some hereditary taint or other. These questions may often be difficult to answer. In Denmark the genetic-hygienic register at the University Institute for Human Genetics, Copenhagen, is often consulted by physicians,—practising physicians, specialists or physicians connected with institutions or hospitals. Then the institute procures information on the family and answers the prognostic question as well as possible. The card-index may furnish information about the persons in question and their families; and at the same time constitute a starting point for medico-genetic studies of the various diseases, which form the necessary basis for all genetic-hygienic activity.

During the period 1939–1952 the Institute for Human Genetics received nearly 8000 inquiries concerning genetic-hygienic problems. These inquiries came from

Physicians connected with hospitals in 1152 cases.

Physicians connected with Mother's Aid Institutions in 5623 cases.

Physicians connected with other institutions in 473 cases.

General practitioners or specialists in 711 cases.

These inquiries were distributed as shown in table 1.

During the first years of this activity the number of such inquiries was only small, but it has been increasing year by year.

These inquiries may be classified according to the diseases which it was feared that expected children might have to suffer from—as shown in table 2 which gives a sample (740 cases) from the year 1949.

Table 1. Genetic-hygienic advice given to physicians by the University Institute for Human Genetics in Copenhagen 1939–52.

Year	Legal abort.	Legal steril.	Legal steril. + abort.	Adoption	Genetic counseling	Total
1939	3	0	0	0	3	6
1940	25	1	0	0	4	30
1941	21	1	0	0	10	32
1942	101	8	0	6	26	141
1943	150	9	1	10	40	210
1944	158	25	15	91	50	339
1945	174	28	34	71	58	365
1946	293	36	37	84	50	500
1947	364	29	51	178	58	680
1948	432	37	48	246	75	838
1949	591	66	68	269	58	1052
1950	702	37	87	197	55	1078
1951	741	62	112	236	69	1220
1952	969	63	110	247	79	1468
Total	4724	402	563	1635	635	7959

Table 2. 740 cases of Genetic counseling classified according to the disease or defect concerned (1949).

	555 cases of induced abortion		132 cases of sterilization		53 cases of genetic-hygienic advice	
	Number	%	Number	%	Number	%
Physical malformations	40	7,2	8	6,1	4	7,5
Deaf-mutism or lesions	9	1,6	4	3,0	2	3,8
Blindness or eye lesions	15	2,7	7	5,3	2	3,8
Skin diseases	3	0,5	2	1,5	1	1,9
Lesions of internal organs	30	5,4	5	3,8	6	11,3
Nervous diseases	108	19,5	9	6,8	15	28,3
Inferioritas, mentally retarded	56	10,1	36	27,3	3	5,7
Feeble-mindedness	45	8,1	23	17,5	6	11,3
Psychosis	73	13,2	11	8,3	2	3,8
Psychopathy	76	13,7	13	9,8	3	5,7
Alcoholism and/or criminality	63	11,4	10	7,6	0	0
Combined eugenic encumbrance	34	6,1	4	3,0	0	0
Various other lesions	3	0,5	0	0	9	16,9
Total	555	100	132	100	53	100

On going through the individual entries in table 2, we find that physical malformations comprise a large number of different defects, e.g. harelip and cleft palate, anencephaly, dysostosis cranio-facialis, osteogenesis imperfecta, chondrodystrophy, craniorhachischisis, splithands and feet, ectrodactylism, congenital amputations, syndactylia and brachydactylia, clubfoot, congenital dislocation of the hip or of other joints, multiple exostoses, malformations of the genitals and several other malformations.

Among the ear lesions that may be the reason for genetic counseling by far the most frequent is the sporadic recessive deaf-mutism, the next are labyrinthine deafness, partial or total, and otosclerosis.

Among the eye-lesions in connection with genetic-hygienic advice we may mention aniridia, retinitis pigmentosa, anophthalmos, microphthalmos, cataract, optic nerve atrophy, microphthalmia, myopia, coloboma iridis, glaucoma, strabismus, nystagmus and ptosis palpebrae.

Among skin diseases the following may be mentioned: ichthyosis cong. and vulgaris, eczema, prurigo *Besnier*, psoriasis, albinismus, and alopecia cong. totalis.

Among the diseases of internal organs the most frequent are diabetes mellitus, hemophilia, thrombopenia, fibropenia and other haemorrhagic diatheses, haemolytic jaundice, congenital cardiac defects, allergic diseases, polyarthritis (rheumatoid arthritis), adiposity, *Graves'* disease, myxoedema and certain types of tumours. Furthermore, erytroblastosis foetalis must be mentioned and cases of toxoplasmosis or rubella infections during pregnancy.

It is very common that questions of genetic-hygienic advice are prompted by the fear that an expected child might suffer from epilepsy. Other nervous diseases are more infrequent. The most common are *Huntington*'s chorea, paralysis agitans, muscular dystrophy, dystrophia myotonica, congenital hyperthropic myotonia, spinal muscular atrophy, infantile progressive muscular atrophy, *Friedreich's* ataxia, nervosismus and neurosis, diffuse cerebral sclerosis, pseudosclerosis, amyotonia, lipothymia, migraine, neurofibromatosis and cerebral palsies.

Among the forms of oligophrenia, the defects mentioned have been mental inferiority (retarded), feeble-mindedness, mongoloid idiocy, amaurotic idiocy, tuberous or diffuse cerebral sclerosis, cerebral atrophy, mb. *Sturge-Weber*, mb. *Niemann-Pick*, mb. *Little* and mb. *Alzheimer*.

Among the psychoses the most common have been schizophrenia, manic-depressive psychosis or endogenous mental depression, compulsion neurosis, anxiety neurosis or other types of neurosis, or related abnormalities such as dyslexia or stuttering.

Another large group includes psychopathy, psychopathic or abnormal personalities, often associated with criminality, alcoholism, asociality, vagrancy, suicide or sexual perversion.

As was to be expected, combined encumbrance was encountered in a certain percentage of the cases, often with several affections—as, for instance, physical malformations together with epilepsy, oligophrenia or other mental lesions. Such cases of combined encumbrance may be very difficult to judge, as no general rule can be laid down concerning them. Every case has to be submitted to individual expert estimation.

In connexion with the genetic counseling a comprehensive empirical material has been collected concerning hereditary diseases and defects, which has been published in Opera ex Domo Biologiae Hereditaria Humanae Hafniensis, Vol. 1–32, 1941–53 and reviewed in *T. Kemp:* "Genetics and Disease", Copenhagen, Edinburgh and London 1951.

It is impossible to predict to how great an extent genetic-hygienic measures prevent hereditary diseases. To judge from theoretical calculations, the negative selection, which can be affected by moderate genetic-hygienic measures, may cause a considerable fall in the incidence of the hereditary diseases in the population. This incidence depends, however, on so many different factors, that the stated calculations are of limited value. On the basis of a genetic-hygienic registration carried through for several generations it will be possible to ascertain whether the hereditary diseases in the population decrease or increase in frequency, and whether the genetic-hygienic measures are employed to suitable extent and in the diseases where they are most needed.

Medical genetics, in connection with the associated genetic-hygienic registration and genetic counseling, are forming the indispensable scientific basis for carrying out the measures necessary to prevent hereditary diseases.

Summary.

Nearly 8000 cases of genetic-hygienic advice, given to physicians during the period 1938–52 by the University Institute for Human Genetics, Copenhagen, are discussed.

The importance of genetic counseling for genetic hygiene is emphasized.

Résumé.

Presque 8000 cas de consultations eugéniques donnés pendant la période de 1938–52 aux médecins par l'Institut d'Hérédité Humaine de l'Université de Copenhague sont passés en revue.

L'importance de la consultation eugénique pour la réalisation des mesures d'hygiène héréditaire est soulignée.

Zusammenfassung.

Etwa 8000 Fälle genetisch-hygienischer Beratung gegeben 1938–52 von dem Universitäts-Institut für menschliche Erblichkeitsforschung in Kopenhagen an Ärzte, sind erwähnt.

Die Bedeutung eugenischer Beratung für die Erbhygiene ist hervorgehoben.

LITERATURE

Hofsten, N. v.: Hered. *37*, 157–265, 1951. – *Kemp, T.:* Arvehygiejne (Genetic Hygiene. With an English summary). Op. dom. Biol. hered. hum. Kbh. *25*, 1951; Genetics and Disease, Copenhagen 1951. – *Lamy, M.:* Précis de Génétique Médicale, Paris 1952. – *Nachtsheim, H.:* Für und wider die Sterilisierung aus eugenischer Indikation, Stuttgart 1952. – *Stern, C.:* Principles of Human Genetics, San Francisco 1949.

Lundström, A.: Acta genet. *4*, 247–257, 1953

From the Department of Orthodontics at the Royal School of Dentistry, Stockholm
(Head: Professor Anders Lundström, Dr. Odont)

AN ANTHROPOLOGICAL COMPARISON OF THE DENTITIONS OF 71 GREEK, 69 TURKISH AND 140 SWEDISH BOYS OF ABOUT 13 YEARS OF AGE

By ANDERS LUNDSTRÖM

The literature would seem to contain few comparable investigations of the dental arches and occlusion for different nationalities. This article is a report of the results of such a comparison between Greek, Turkish and Swedish boys.

The Greek boys were 71 unselected pupils of the Pancyprion Secondary School in Nicosia, Cyprus, while the 69 Turkish boys were at the Turkish Lycée in the same town. The pupils of the age-group studied (Table 1) came mainly from rather well-situated families, since the secondary school education is not compulsory.

It appears that for religious reasons there is no intermarriage between the Greek and Turkish populations in Cyprus, so that the children studied would in fact come from distinct population groups.

The 140 Swedish children were in the last class of a Stockholm elementary school.

It is clear that the material studied is not representative of the Greek, Turkish and Swedish populations, but this does not detract unduly from the value of a comparison between the groups.

Table 1. Mean and standard deviation of the age (in years) for Swedish, Greek and Turkish boys.

	Swedes	Greeks	Turks
Number	140	71	69
Mean	13.61	13.55	13.56
Standard deviation . .	0.38	0.40	0.85

The first question to be decided is whether these population groups can be considered as representing different races. When the term "race" is used one often thinks of an original state of "pure races", which have subsequently inter-mixed to a varying degree. There is, however, no proof that pure races have ever existed (*Dahlberg* [1952]). If the definition "race" given by *Dahlberg* [1941, 1952] is used it does seem justifiable to count these population groups as different races. According to this definition a race is "a geographically or socially determined isolate, or sub-division of an isolate, that is genetically distinct from other isolates". From the point of view of the dentition this question is of interest in so far as it is possible to establish genetically distinct dental characteristics for different races.

The Swedish material was collected in 1939 and 1940, and the Greek and Turkish in 1952, simultaneously with investigations that were carried out in Cyprus by the Swedish Expedition for Archeological Anthropology under the direction of Professor *C. H. Hjortsjö*, Lund.

Properties studied.

The dental properties to which attention was given were:

1. tooth width
2. dental arch dimensions in the upper jaw
3. height of palate
4. relative spacing of the teeth in the upper jaw
5. inclination of the upper incisors
6. mutual relationship of the dental arches

For the Greek and Turkish groups the following registrations and measurements were performed:

7. number of decayed teeth
8. head length
9. head width
10. bizygomatic width

In addition to the material mentioned, reference was made to the journals of the school dental service in respect of a further 63 Stockholm elementary school boys of 13 years of age.

Methods of registration.

The registration of the various dental characteristics was largely performed in accordance with the measuring procedures previously described by the present author [1948]. A somewhat simplified method was used however for the tooth spacing in the case of the Greeks and Turks. In cases where neither overcrowding nor overspacing was present at the second premolars, the space difference (= the difference between the arch perimeter measured sectionally for each pair of teeth and the sum of the individual tooth widths) was calculated for the anterior part of the arch from the contacts between the first and second *premolars* instead of from the contact between the first and second *molars*. This modification of the method is unlikely to have any appreciable influence on the measurement values.

Another difference in the measuring procedure between the Swedish boys and the Greeks and Turks was that for the Cypriots all measurements were performed directly in the mouth, while for the Swede this was not so in the case of the first premolar widths, the arch dimensions, the palatal height, the tooth-spacing, the inclination of the upper incisors and the mutual relationship between the arches. For the Swedish boys these properties were recorded on plaster casts of the jaws prepared from hydrocolloid impressions. This difference in measurement procedure is not wholly without significance, as the present author has shown that the tooth widths on the cast are about 1 % larger than those given by direct measurement (1943).

For the measurement of the length of the arch, the palatal height and the inclination of the lateral incisors use was made of a special instrument which has been described elsewhere (1946, 1948). In the case of the Greeks and Turks the length of the arch and the inclination of the upper incisors were recorded with the same instrument setting, in contrast to the earlier described method when separate settings were used for these two measurements.

The instrument was mounted in the way described for measuring the incisor inclination, after which the length was measured from the line connecting the first molars to the anterior of the glide. Checks showed that the figures so obtained were on an average 3.0 mm too

great; accordingly, the averages for the Greeks and Turks have been reduced by this amount in Table 3.

The anthropological measurements of the cranial length and width and the bizygomatic width were carried out by the procedure described by *Martin*.

The caries registration was undertaken with the aid of mirror and probe (sharply pointed). Well-defined carious defects have been distinguished from less definite defects where the probe certainly caught up in the fissure system but where it was still uncertain whether a true carious lesion was present.

Results.

The results are given in Tables 1–6. The most clearly evident difference is found in association with the caries incidence (Table 5).

Table 2. Comparison of Swedish, Greek and Turkish boys concerning mesio-distal widths of incisors, canines and first premolars. N = number of cases, M$\pm$ ε(M) = mean $\pm$ stand. error of mean, σ = standard deviation. N = number of cases.

Tooth	Swedes			Greeks			Turks		
	N	M $\pm$ ε(M)	σ	N	M $\pm$ ε(M)	σ	N	M $\pm$ ε(M)	σ
				upper jaw					
I_1 right	170	8.66$\pm$0.04	0.53	67	8.64$\pm$0.07	0.55	69	8.71$\pm$0.06	0.51
left	176	8.69$\pm$0.04	0.52	69	8.70$\pm$0.07	0.59	69	8.76$\pm$0.06	0.52
I_2 right	175	6.64$\pm$0.04	0.58	70	6.62$\pm$0.06	0.52	68	6.66$\pm$0.06	0.50
left	180	6.72$\pm$0.04	0.52	70	6.75$\pm$0.06	0.52	68	6.78$\pm$0.06	0.49
C right	173	7.80$\pm$0.03	0.43	60	7.79$\pm$0.05	0.41	56	7.83$\pm$0.06	0.44
left	170	7.84$\pm$0.03	0.41	60	7.78$\pm$0.06	0.44	61	7.83$\pm$0.05	0.40
P_1 right	111	7.15$\pm$0.03	0.36	70	6.92$\pm$0.05	0.43	66	6.85$\pm$0.04	0.33
left	110	7.17$\pm$0.04	0.39	69	6.92$\pm$0.05	0.40	66	6.92$\pm$0.04	0.32
				lower jaw					
I_1 right	178	5.33$\pm$0.02	0.32	71	5.47$\pm$0.04	0.32	69	5.45$\pm$0.04	0.33
left	178	5.37$\pm$0.02	0.31	70	5.49$\pm$0.04	0.30	68	5.48$\pm$0.04	0.35
I_2 right	179	5.93$\pm$0.03	0.37	71	6.04$\pm$0.04	0.37	68	6.04$\pm$0.04	0.35
left	182	5.96$\pm$0.03	0.36	70	6.07$\pm$0.05	0.40	68	6.01$\pm$0.04	0.37
C right	184	6.91$\pm$0.03	0.38	69	6.80$\pm$0.05	0.42	64	6.85$\pm$0.04	0.36
left	181	6.92$\pm$0.03	0.38	67	6.83$\pm$0.05	0.39	67	6.91$\pm$0.04	0.36
P_1 right	128	7.21$\pm$0.03	0.38	70	6.93$\pm$0.05	0.45	65	6.92$\pm$0.05	0.44
left	125	7.19$\pm$0.03	0.37	70	6.97$\pm$0.06	0.49	69	6.96$\pm$0.05	0.38

[1] The Swedish figures for the first premolars are not strictly comparable with the Greek and Turkish figures, as the former were obtained from measurements on plaster models and the latter from direct measurements.

Large and obviously significant differences exist between the Swedish children on the one hand and the Greeks and Turks on the other. A closer analysis of the differences in diet between the various populations should prove rewarding, but would clearly demand very thorough and accurate study.

A factor that should be observed in this connection is the fluorine content of the drinking water. In some regions of Cyprus (some villages in the Larnaca district) there is general mottled enamel. An examination of the drinking water in two of these villages – Arsos and Melousha – (made by the National Institute of Public Health in Stockholm) showed 3 samples to contain 1.4, 1.6 and 2.0 mg fluorine per litre. These figures clearly exceed the boundary value, of 1 mg per litre above which mottled enamel is generally reckoned to occur. The Greek and Turkish boys presented mottled enamel in 7 and 2 cases respectively.

No difference in the dental arches or occlusion could be established between children with mottled enamel and others, although one of the two Greek boys with this disorder also presented exceptional crowding of the teeth.

The villagers of Melousha (fluorine content = 1.6 mg/litre) declared that greying of the hair was often observed to coincide with early loss of teeth. There was unfortunately no means of checking this. Even if confirmation had been possible no conclusions could have been drawn solely on that basis, since the village may be regarded as an isolate where quite a high incidence of certain diseases or anomalies might obtain without there being any direct causal connection.

From the anthropological aspect prominence might be given

Table 3. Comparison of Swedish, Greek and Turkish boys concerning: width of the upper dental arch at the first premolars (= B_1) and at the first molars (= B_2), length of the upper dental arch from a line connecting the first molars to the central incisors (= L), height of the palatal vault at the first molars (H).

Measurement	Swedes			Greeks			Turks		
	N	M ± s (M)	σ	N	M ± (s M)	σ	N	M ± s (M)	σ
B_1	112	36.7±0.2	2.4	69	36.4±0.29	2.4	65	36.9±0.32	2.6
B_2	111	47.4±0.2	2.5	68	48.2±0.35	2.9	67	49.0±0.32	2.6
L	112	31.6±0.2	2.3	67	31.5±0.26	2.1	66	31.6±0.25	2.0
H	128	17.2±0.14	1.6	70	16.1±0.24	2.0	67	15.5±0.21	1.7

in particular to the width-length index of the head. The Greeks have clearly a somewhat more dolichocephalous form than the Turks (diff. = 3.06 ± 0.68). This difference is mainly due to the length of the heads, while the widths would seem to be rather similar (Table 6).

There is a remarkable similarity between the Swedes, Greeks and Turks in respect of the mean values of the various dentition characteristics. The Turkish boys, however, show a somewhat wider maxillary arch at the first molars than the Swedish boys (diff. = 1.6 ± 0.38 mm: see Table 3). The Swedish boys have, furthermore, somewhat higher palatal vaults than the Greek (diff. = 1.0 ± 0.27 mm) and the Turkish boys (diff. = 1.7 ± 0.25 mm).

Concerning the arch-width at the first premolars and the first molars, *Seipel* [1946] has presented somewhat lower figures for Swedish 13-year old boys than the author's and the differences are significant. It is of course conceivable that systematic differences in measurement have contributed to this result. Such errors might be suspected as fissure fillings, might quite often render the measurement points difficult to determine with certainty in Swedish present day material. The fact that the author measured on models while Seipel measured directly might also have some significance.

One factor, that must also be taken into consideration in this connection is the loss of certain parts of the material through extractions of teeth. This loss is less in *Seipel*'s material (about 12 per cent.) than in the author's (about 20 per cent.). It is probably safe to assume that the majority of extractions were due to deep carious

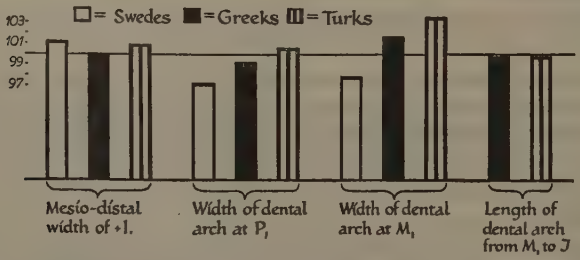


Fig. 1. Comparison of Swedish, Greek and Turkish boys in respect of some investigated characteristics. The horizontal line represents the present author's means for Swedish boys. The staples represent the percentage deviation from these figures of a) *Seipel*'s means for Swedish boys, b) the author's means for Greek boys, c) the author's means for Turkish boys.

lesions without correlation to the size of the dental arch. In a small proportion of cases, however, the extractions were performed as treatment for crowding in narrow or short jaws. This factor might have given rise to a small displacement of the means in the direction of too wide arches in comparison with the true mean of the population. The higher percentage loss in the author's material might thus have given a somewhat higher mean than *Seipel*'s.

Summing up, it is obviously difficult to say whether the author's or *Seipel*'s figures are preferable for purposes of comparison. In Fig. 1 both materials are therefore presented.

With regard to the space difference which provides an expression of the variation from marked crowding to marked overspacing, it is seen that the regular upper dental arch is the average for all three groups of the material. (Table 4 and Fig. 2.) (With the method used the perfectly regular arch has a value of about + 1 mm – see *Lundström* [1951].) The differences between the materials as expressed by the standard deviation, are not great. It is, however, probable that the dispersion is somewhat less for the Turkish boys than for the Swedes and Greeks (diff. = 1.0 ± 0.34 mm and 0.9 ± 0.39 mm resp.). When assessing the Swedish figures it must again be borne in mind that the values apply to boys with no extracted teeth. In this selection 28 of the 139 boys were excluded. To the extent that this selection was made on account of extraction due to deep caries lesions (especially permanent first molars) there was probably no effect on the distribution of the space difference.

Since, however, some of the extractions in the Swedish material were undertaken as treatment for more pronounced crowding it is not possible to neglect a certain influence on both the mean, and the

Table 4. Comparison of Swedish, Greek and Turkish boys concerning: inclination of upper central incisors to occlusal plane (incl. +1+), horizontal (H.o.b.) and vertical (V.o.b.) overbite, space-difference in the upper jaw (Sp.-diff.).

Measurement	Swedes			Greeks			Turks		
	N	M ± s(M)	σ	N	M ± s(M)	σ	N	M ± s(M)	σ
Incl.+1+	140	81°.8±0°.42	5°.0	69	83°.6±0°.51	4°.2	68	82°.2±0°.62	5°.1
H.o.b.[1]	138	3.0±0.14	1.6	69	3.5±0.18	1.5	69	3.1±0.19	1.6
V.o.b.[1]	138	3.0±0.12	1.4	69	3.7±0.19	1.6	69	3.3±0.19	1.6
Sp.-diff.[1]	111	+0.8±0.35	3.7	67	+1.0±0.44	3.6	67	+0.7±0.33	2.7

[1] in mm.

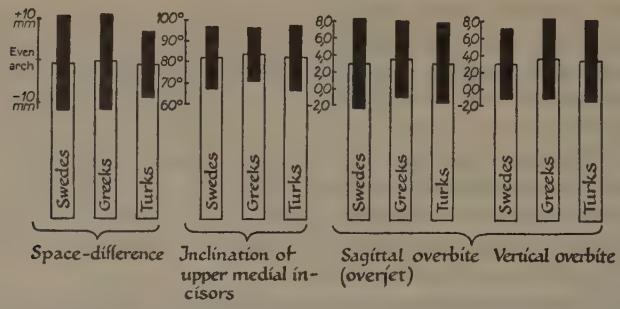


Fig. 2. Comparison of Swedish, Greek and Turkish boys in respect of some investigated characteristics. The white staples give the means and the black staples the variation range around these means, expressed as ± 3 times the standard deviation.

Table 5. Comparison of 63 Swedish, 71 Greek and 69 Turkish boys concerning number (N) and percentage of carious incisors, first premolars and first molars. The Swedish figures comprise teeth with fillings, the Greek and Turkish figures teeth with fillings or carious cavities. In Greeks and Turks figures are given for teeth with definite caries and for teeth with uncertain caries (within brackets).

Tooth	Swedes		Greeks		Turks	
	N	Percent. carious	N	Percent. carious	N	Percent. carious
			upper jaw			
I_1 right	29	46.0±6.3	4	5.6±2.7	1	1.4±1.4
left	31	49.2±6.3	5	7.0±3.0	1	1.4±1.4
I_2 right	30	47.6±6.3	4 (1)	5.6±2.7	1	1.4±1.4
left	31	49.2±6.3	3 (1)	4.2±2.4	2	2.9±2.0
P_1 right	15	23.8±5.4	1 (1)	1.4±1.4	1	1.4±1.4
left	13	20.6±5.1	1 (1)	1.4±1.4	1	1.4±1.4
M_1 right	57	90.5±3.7	6 (12)	8.5±3.3	3 (3)	4.3±2.4
left	56	88.9±4.0	10 (10)	14.1±4.1	4 (3)	5.8±2.8
			lower jaw			
I_1 right	13	20.6±5.1	0	0	0	0
left	14	22.2±5.2	0	0	0	0
I_2 right	16	25.4±5.5	0	0	0	0
left	13	20.6±5.1	0	0	0	0
P_1 right	6	9.5±3.7	1 (1)	1.4±1.4	0	0
left	5	7.9±3.4	0 (1)	0	0	0
M_1 right	54	85.7±4.4	9 (14)	12.7±4.0	5 (8)	7.2±3.1
left	57	90.5±3.7	8 (14)	11.3±3.8	7 (13)	10.1±3.6

dispersion. It would seem likely, however, that this effect is small if the relatively small number of cases of this kind is taken into account (*Lundström* [1951]).

The mean values and the dispersion for overjet and overbite and for the inclination of the upper central incisors are much the same for the three groups (Table 5). For overjet and overbite *Seipel* [1946] has presented figures for Swedish boys differing somewhat from the author's. As this discrepancy might be due to the fact that the measurements were made by different persons, the author's values would seem preferable for comparison.

The close similarity in the tooth position and occlusion between the three groups, and the very great differences in the caries incidence between, on the one hand, the Swedish boys and, on the other, the Greeks and Turks does not support the existence of a definite parallel trend between an increase in caries and malocclusion of the teeth.

It seems likely that the dentitional variations presented by this material are first and foremost inherent basic characteristics of the population groups. Such a supposition is in fair accord with the apparent importance of hereditary and internal environmental factors to these variations (see for example, *Lundström* [1948]).

Table 6. Comparison of 71 Greek and 69 Turkish boys concerning cephalic length and width, cephalic index and width of face (bizygomatical width).

Measurement	Greeks		Turks	
	M ± ε (M)	σ	M ± ε (M)	σ
Cephalic length . .	183.39±0.85	7.2	178.38±0.65	5.4
Cephalic width . .	148.01±0.72	6.0	149.51±0.61	5.1
Cephalic index . .	80.80±0.53	4.5	83.86±0.42	3.5
Bizyg. width . . .	129.15±0.72	6.1	129.29±0.60	5.0

For their kind assistance and help during my school investigations in Nicosia, I wish to express my gratitude to teachers and staff members of the Pancyprion Secondary School (Headmaster: Dr. *C. Spyridakis*) and of the Turkish Lycée (Headmaster: Mr. *Yavouz Connolou*, master in charge of the junior school: Mr. *Hasan Tahsin*).

The statistical treatment of the material has been partly carried out at the State Institute of Human Genetics and Race Biology in Uppsala and it is a pleasure for me to thank Professor *Gunnar Dahlberg* for this valuable help.

Analyses of the fluorine content in proofs of drinking water were made at the National Institute of Public Health in Stockholm and my thanks for this are due to the Director, Professor *Ernst Abramson*.

I also wish to thank the Director of the Public Dental Service of Stockholm, Dr. *H. Berggren*, who made available journals of the school dental service.

At the same time, I would like to acknowledge receipt of financial assistance from the Swedish Medical Research Council.

Summary.

A comparison between three groups from different populations (Swedish, Greek and Turkish boys) showed rather close similarity between the groups as regards tooth-size, tooth-position and occlusion.

There was, however, a very great difference between, on the one hand, the Swedish boys and, on the other, the Greeks and Turks with much higher caries incidence for the former than for the latter.

The findings do not support the existence of a definite parallel trend between an increase in caries and malocclusion of the teeth. This is in fair accord with the assumed significance of hereditary and internal environmental factors to malocclusion of the teeth, while such factors as caries and early loss of deciduous teeth seem to be of secondary importance from an aetiological point of view.

Figures were given of head length, head width and bizygomatic width for the Greek and Turkish boys. The Greeks have somewhat more dolichocephalous heads than the Turks, and the difference is mainly due to the length of the heads while the widths would seem to be rather similar.

Résumé.

Une comparaison entre trois groupes de populations différentes (des garçons suédois, grecs et turcs) a montré une assez bonne concordance entre les groupes en ce qui concerne la grandeur, la position et l'occlusion des dents.

Pourtant il y avait une très grande différence entre, d'un côté, les garçons suédois, et de l'autre, les garçons grecs et turcs quant à la fréquence des caries beaucoup plus élevées chez les premiers que chez les seconds.

Ces constatations ne parlent pas d'un parallélisme entre l'augmentation de la carie et la « malocclusion » des dents. Ceci concorde avec l'importance supposée des facteurs héréditaires et des facteurs du milieu interne pour la « malocclusion » des dents, tandis que des facteurs tels que la carie et la perte prématurée des dents de lait semblent être d'une importance secondaire du point de vue étiologique.

Sur les garçons grecs et turcs on prenait des mesures de la longueur et de la largeur de la tête ainsi que de la largeur du visage. Les grecs sont un peu plus dolichocéphales que les turques. La différence dépend surtout de la longueur de la tête, la largeur paraissant être à peu près identique dans les différents cas.

Zusammenfassung.

Ein Vergleich zwischen drei ungleichen Bevölkerungsgruppen (schwedische, griechische und türkische Jungen) ergab eine ziemlich große Übereinstimmung zwischen diesen betreffend Zahngröße, Zahnstellung und Okklusion.

Doch bestand ein sehr großer Unterschied zwischen schwedischen Jungen einerseits und griechischen und türkischen anderseits im Hinblick darauf, daß die Kariesfrequenz bei ersteren bedeutend höher lag als bei letzteren.

Das Ergebnis liefert keinen Beweis für die Existenz eines definitiv parallelen

Trends zwischen einem Anwachsen der Karies und der Malokklusion bei Zähnen. Dieses steht in gutem Einklang mit der Annahme einer Bedeutung erblicher und interner Milieufaktoren für die Malokklusion bei Zähnen, während solche Faktoren wie Karies und zeitiger Verlust der Milchzähne vom etiologischen Gesichtspunkt von sekundärer Bedeutung zu sein scheinen.

Die Ziffern sind für Kopflänge, Kopfbreite und bizygomatische Breite für griechische und türkische Jungen gegeben. Die Griechen haben etwas dolikozefalischere Köpfe als die Türken. Der Unterschied beruht hauptsächlich auf der Kopflänge, während die Breite ziemlich übereinzustimmen scheint.

LITERATURE

Dahlberg, G. and *S. Wahlund:* "The Race Biology of the Swedish Lapps", Part II, Uppsala 1941. – *Dahlberg, G.:* Acta genet. 3, 1952. – *Lundström, A.:* Svensk tandläk. Tidskr. *36*, 1943; «Vilka förändringar ha inträffat ifråga om käkarnas storlek och occlusionsförhållanden under historisk tid?» Svensk tandläk. Tidskr. *39*, 1946; "Tooth-size and occlusion in twins". S. Karger, Basle and New York, 1948; "Aetiology of crowding of the teeth". Trans. europ. orthod. Soc. 1951. – *Martin, R.:* «Lehrbuch der Anthropologie», Jena, 1928. – *Seipel, C. M.:* "Variation of tooth-position". Svensk tandläk. Tidskr. *39*, Supplement, 1946.

Penrose, L. S.: Acta genet. *4*, 257–265, 1953

Galton Laboratory, University College, London

THE GENETICAL BACKGROUND OF COMMON DISEASES

By L. S. PENROSE

Introduction

A direct method of genetical investigation of common traits is to establish the fact of familial concentration by comparing the relatives of propositi, or index cases, with the general population. A raised incidence in relatives is then plausibly interpreted as indicating that the trait in question has some hereditary background. When the difference between the familial and population incidence is found to be large, say tenfold or more, the result is so striking that the prob-

ability of genetical causation appears to be high even in data compiled with little attention to detail. The surveys, carried out by *Luxenburger* [1936] to demonstrate the familial concentration of various kinds of mental illnesses, are cases in point.

However, in some conditions, the observed difference between the familial and population figures is relatively small so that, to ascertain its extent or even to find whether or not it really exists, great care must be exercised in the collection and analysis of the material. For example, data have to be compiled in such a way that ages and sexes can be separately classified. A good example of the precise application of this method is the study of peptic ulcer made by *Doll* and *Buch* [1950]. These investigators first set up a control based upon the observed incidence of the trait in the general population at each age, specifying the sexes separately. They were thus able to show convincingly that the risk of contracting the disease was about four times as great for parents and twice as great for sibs as for members of the population at large. Using a somewhat similar method, figures have been obtained by *Stecher* [1941] showing that, for females, the risk of developing *Heberden*'s nodes on the fingers was doubled in close relatives of affected cases. In the survey of mammary cancer by *Penrose*, *MacKenzie* and *Karn* [1948] the incidence was shown to be doubled in parents and trebled in sibs of affected cases. A study of the incidence of acute rheumatism in the sibs of affected propositi by *Roberts* [1951] indicated that the familial incidence was not as much as twice that in the general population. In this and most other surveys where sibships are each counted once, the crude incidence in sibs underrates the true value, which lies between the observed value and twice the observed value. When the true familial incidence in sibs is very low it closely approaches twice the observed value. Thus, in *Doll* and *Buch*'s data, the incidence in parents and sibs probably can be regarded as about equal.

In all these examples, the crucial value, the ratio of familial to population incidence, which can be called K, is small, varying between about $1\frac{1}{2}$ and $4\frac{1}{2}$. It is the purpose of this paper to examine some possible interpretations of these low K-values; they do not necessarily indicate the relative absence of genetical determination, as often supposed. Actually, a low value of K indicates that, if the cause of the trait is a gene, this gene must be very common in the population studied. However, if the value of K is actually indistinguishable from unity, it is reasonable to suppose that genetical

differences play no appreciable part in causation in the population in question.

The Standard Formulae.

There are two simple traditional ways of interpreting familial concentration. The observed trait may be due to an autosomal recessive or an autosomal dominant gene. A third interpretation is that the gene is intermediate and, in the present instance, it will be assumed that this means that a trait is manifested in all cases in the homozygote but only in one half of those who are heterozygous. The manifestation of the trait in these heterozygotes could be determined by other genes but, in the present instance, it will be considered to be a threshold phenomenon determined non-genetically. The values of K which cover the range already mentioned are given in the left hand column of Table 1, and the corresponding frequencies of a gene interpreted as being the cause are set out.

Marked differences between parent-child and sib-sib relationships are shown in recessive traits. However, when the K-values are very low, indicating very common genes, the distinction between recessive and dominant hypotheses by this means may be difficult unless data are extensive. The intermediate hypothesis as defined here gives rise to values of q which are the same for both types of relationship and they do not differ much from those consistent with the single dominant gene hypothesis.

Table 1. Single gene frequency q, for different values of K, i.e. incidence in relative as compared with population incidence.

K	Recessive Gene		Intermediate Gene		Dominant Gene	
	Parent or child	Sib	Parent or child	Sib	Parent or child	Sib
1.0	1.000	1.000	1.000	1.000	1.000	1.000
1.5	0.667	0.690	0.333	0.333	0.262	0.281
2.0	0.500	0.547	0.200	0.200	0.170	0.178
3.0	0.333	0.406	0.111	0.111	0.100	0.103
5.0	0.200	0.288	0.059	0.059	0.056	0.056
10.0	0.100	0.188	0.027	0.027	0.026	0.026
Indefinitely large	1/K	$1/2K^{½}$	1/4K	1/4K	1/4K	1/4K
General formula	$qK=1$	$q(2K^{½}-1)=1$	$q(4K-3)=1$	$q(4K-3)=1$	$(2-q)^2K=1+q-q^2$	$4q(2-q)^2K=4+5q-6q^3+q^3$

Trait and Disposition.

The great majority of traits or diseases thought to be strongly influenced by genes are also strongly influenced by environment. That is to say, quite apart from the effects of dominance, recessiveness and intermediacy, agreement between phenotype and genotype is not perfect. The trait for which there exists a genetical susceptibility or predisposition may not always be manifested, indeed it may be so only very rarely. In addition to this, the trait might sometimes occur in the absence of the predisposing genetical background. Such acquired cases, fresh mutations or phenocopies, are generally supposed to be very infrequent as compared with those occurring on the genetical basis; yet they must not be neglected in human data because cases are recorded when the specific outward effect is produced whatever the cause. The alternatives are set out in Table 2. There are four different classes, containing x, y, z and w people respectively. The total number affected is $(x + y)$ and the number genetically predisposed is $(x + z)$.

Table 2. Relationship between trait and hereditary disposition.

Class	Number of people	Description
(i)	x	Affected and also genetically predisposed (hereditary, endogenous or primary cases)
(ii)	y	Affected but not genetically predisposed (acquired, environmental, exogenous or secondary cases)
(iii)	z	Unaffected although genetically predisposed (latent, susceptible)
(iv)	w	Unaffected and not genetically predisposed

The proportion of predisposed cases, M, who are actually affected and show the trait, is $(x)/(x + z)$; this is sometimes termed penetrance though it is more correctly termed manifestation since, in human genetics, the distinction between penetrance and expressivity is invalid. The importance of the value of $(x)/(x + z)$ in the present discussion is that, provided y is a negligible quantity, x measures the incidence of the trait. Furthermore, when y is zero, the estimate of gene frequency in Table 1 is independent of the manifestation. The disposition frequency, which is the most significant factor for purposes of estimating eugenic prognosis or the value of preventive

measures, can be ascertained directly from the gene frequency after the mode of inheritance, recessive, intermediate or dominant, has been determined.

In the case of the hypothesis of intermediacy, for example, we can suppose that, over and above the indeterminancy of the threshold effect in the heterozygote, there is an imperfect degree of manifestation. The supposition then would be that genotype AA would be always unaffected (given that $y = 0$), genotype Aa would be affected in $M/2$ instances and aa would be affected in M instances. The relationship between parent and child (or between sib and sib) would then be as shown in Table 3. The gene frequency, in the intermediate case as here defined, is equal to the frequency of predisposed people, the incidence of those actually showing the trait being Mq.

Table 3. Correlation between parent and child for an incompletely manifested intermediate trait.

Child / Parent		Trait: Absent	Trait: Present	Total
Trait	Absent	$1-2Mq+\frac{1}{4}M^2q(1+3q)$	$Mq-\frac{1}{4}M^2q(1+3q)$	$1-Mq$
	Present	$Mq-\frac{1}{4}M^2q(1+3q)$	$\frac{1}{4}M^2q(1+3q)$	Mq
Total		$1-Mq$	Mq	1

M = manifestation; q = gene frequency = predisposition frequency = F; K = $(1+3q)/4q$, independent of M.

The value of K, which is the proportion affected among relatives of affected propositi, i.e. $\frac{1}{4}M\ (1+3q)$, divided by the population incidence, i.e. Mq, is equal to $(1+3q)/4q$ and is independent of manifestation.

These considerations make it convenient to express Table 1 in a different, equivalent, form, as shown in Table 4, where the predisposition frequencies, F, are given instead of gene frequencies.

Multiple Gene Effects.

The intermediate gene hypothesis is similar in certain respects to the hypothesis of multiple additive genes. Gene frequency has no obvious meaning for characters due to multiple genes, except as a rough average value, but frequency of the predisposition is still a very significant quantity. According to the definition of intermediacy used here, a trait determined in this way would give parent-child and sib-sib correlations of 0.25 for all gene frequencies. Multiple alternative additive genes give rise to parent-child and sib-sib correlations of

Table 4. Predisposition frequency, F, for different values of K,

K	Recessive Gene		Intermediate Gene
	Parent or child	Sib	Parent or child
1.0	1.000	1.000	1 000
1.5	0.444	0.476	0 333
2.0	0.250	0.299	0.200
3.0	0.111	0.165	0.111
5.0	0.040	0.083	0.059
10.0	0.010	0.035	0.027
Indefinitely large	$1/K^2$	1/4K	1/4K
General formula	$FK^2=1$	$F(2K^{½}-1)^2=1$	$F(4K-3)=1$

0.5, when the character which they determine has a continuous distribution, but an apparent reduction is produced by the all-or-none classification of affected and unaffected people. Values of K corresponding to different degrees of F, on the assumption that the genetical background is continuous and the genotypical correlation 0.5, can be calculated from *Pearson*'s [1914] tables for tetrachoric functions. Within the range $K = 1.5$ to $K = 10$, the values of F obtained in this way differ little from those for the single intermediate gene, as shown in Table 5. The value of the observed crude correlation coefficient gradually diminishes as K increases but, at the value of $K = 3.2$ it is equal to 0.25.

The critical distinction between the hypothesis of single dominant or intermediate gene and multiple genes cannot be made unless other data than K-values are available. One kind of data concerns monozygotic twins. According to the present interpretation of the action

Table 5. Hypotheses of single intermediate gene and multiple additive alternative genes compared; values of F for different values of K in parents or sibs.

K	Single Intermediate Gene predisposition frequency F	Multiple Additive Genes predisposition frequency, F
1.0	1.000	1.000
1.5	0.333	0.384
2.0	0.200	0.239
3.0	0.111	0.115
5.0	0.059	0.048
10.0	0.027	0.015

i. e. incidence in relative as compared with population incidence.

Intermediate Gene	Dominant Gene	
Sib	Parent or child	Sib
1.000	1.000	1.000
0.333	0.455	0.483
0.200	0.311	0.324
0.111	0.190	0.195
0.059	0.109	0.109
0.027	0.052	0.052
1/4K	1/2K	1/2K
F(4K—3) =1	FK[1+(1—F)½] =F+(1—F)½	4FK[1+(1—F)½] =1+3F+3(1—F)½+F(1—F)½

of an intermediate gene, the expected correlation between monozygotic twins would be exactly 0.5; and the incidence of the trait in the monozygotic twin of a propositus would be $F = 1/(2K - 1)$.

Assuming that the threshold effect producing the trait is a sharp one, when multiple additive genes are involved, this correlation would be higher than 0.5 in monozygotic twins and the corresponding F-value would approach unity. On the other hand, the more genes which are involved in the manifestation of a trait, the lower will be the familial incidence. Thus a relatively strong likeness between monozygotic twin pairs, together with a weak likeness between dizygotic pairs or ordinary sibs, is suggestive of multiple gene causation with a sharp threshold controlling manifestation. In making any practical use of this suggestion, however, the effects of incomplete manifestation due to environmental causes, both on twin pairs and other kinds of relatives, would have to be taken into account. The consequent reduction of all these incidence rates if M is less than unity may not always be proportional. But, as a general rule, if the incidence of a trait in monozygotic twins of propositi manifesting a trait is more than four times that in dizygotic twins, multiple alternative gene determination is probable.

Examples.

In the survey of the familial incidence of mammary cancer mortality (*Penrose*, *MacKenzie* and *Karn* [1948]), for parents of affected propositi, the incidence was 25 against an expected value of 11.3 based upon the general population, leading to a K-value of 2.21.

For sibs, the incidence was 24 against 7.05, leading to a *K*-value of 3.38. We might reasonably suppose that the lower incidence in parents than in sibs indicates recessivity of a single causal gene and that such a gene could have a frequency of about 0.4. This corresponds to a predisposition frequency of 0.16. The high prevalence of the disposition agrees with the observation that the condition is responsible for some three per cent of the mortality of all females. In males, the manifestation is very low, only about one-hundredth of the female rate, although it could be assumed that the disposition is equally frequent in the two sexes. Such an explanation is not inconsistent with the view that the manifestation of the disease is also determined by environment or by very prevalent cytoplasmic agents transmitted through the parents, especially the mother. The main point to be emphasized, however, is not the possibility of single gene causation but the very great prevalence, which any recessive gene must have, to account for the observed familial incidence.

In the material of *Doll* and *Buch* [1950] on peptic ulcer the *K*-values are consistent with causation by a single dominant gene with frequency rather less than 10 per cent, or with an intermediate or multiple gene hypothesis with predisposition incidence of the same order of magnitude. Since the incidence in the adult control population was over 5 per cent in males and only 1 per cent in females, the gene may be manifested almost fully in males in some age groups though always incompletely in females. There is no hint of sex linkage in either this or the material on mammary cancer, so that it is reasonable to think in terms of autosomal genes as causal factors in both examples with very marked degrees of sex limitation. Further data were supplied by the same writers on the probable genetical distinction between gastric and duodenal sites. Since the *K*-value for ulcers at the duodenal site is higher than that for the gastric site, its predisposing genic background should be the less common of the two.

These analyses emphasize that the degree of familial incidence is not in itself an indication of degree of hereditary causation: it is an index of frequency of the hereditary predisposition in the population which is being studied. When the genetical basis of a trait is very common, however, as in the examples just given, the genetical causation is of less practical interest than environmental causes. In the example of acute rheumatic fever, the value of *K* being barely 1.5 is compatible with a frequency of at least one third for an intermediate gene; the susceptibility must, on any single gene hypothesis, be

present in nearly half the population. There might be, on the other hand, many alternative additive genes all with about the same frequency but giving rise to a much less common susceptibility. A relatively high value of K in the monozygotic twin of an affected propositus, as compared with K for sibs, would point towards multiple genes rather than to a single gene pair. In fact, there is some evidence on this point compiled by *Stevenson* and *Cheeseman* [1953]. Numerous instances of monozygotic twin pairs, both affected, have been reported but few of such pairs with only one affected. The incidence in monozygotic twins is certain to be more than 50 per cent whereas the corrected incidence in sibs or in dizygotic twins of affected propositi is less than 10 per cent. Many alternative common genes are therefore indicated as probable part causes.

Summary.

The principles of genetical analysis of very common traits by the method of comparing familial with population incidence are stated. The effects of imperfect correspondence between genotype and phenotype are discussed. Autosomal gene frequencies and predisposition frequencies, agreeing with various hypotheses, are tabulated and some examples of their application described.

Résumé.

L'auteur expose les principes d'une analyse génétique de caractères très fréquents en comparant leur presence dans les familles avec leur présence dans la population. Les effets d'un accord imparfait entre le génotype et le phénotype sont discutés. La fréquence des gènes autosomiques et la fréquence de prédisposition, s'accordant avec des hypothèses différentes, ont été groupées sous forme de tables, et leur application est illustrée par quelques exemples.

Zusammenfassung.

Der Verfasser legt die Prinzipien einer genetischen Analyse sehr gewöhnlicher Eigenschaften dar, indem er Vorkommen in Familien mit solchem innerhalb der Bevölkerung vergleicht. Die Auswirkungen unvollständiger Übereinstimmung zwischen Genotypus und Phänotypus wurden erörtert. Autosomatische Genfrequenzen und Prädispositionsfrequenz, die mit verschiedenen Hypothesen übereinstimmten, wurden in Tabellenform gebracht und einige Beispiele für deren Anwendung beschrieben.

REFERENCES

Doll, R. and *J. Buch:* Ann. Eugen., Lond. *15*, 135, 1950. – *Luxenburger, H.:* Zbl. ges. Neurol. Psychiat. *84*, 585, 1936. – *Pearson, K.:* Tables for statisticians and biometricians. Cambridge University Press, London 1914. – *Penrose, L. S., H. MacKenzie* and *M. N. Karn:* Ann. Eugen., Lond. *14*, 234, 1948. – *Roberts, L.:* Med. Offr. *86*, 159, 1951. – *Stecher, R. M.:* Amer. J. med. Sci. *201*, 801, 1941. – *Stevenson, A. C.* and *E. A. Cheeseman:* Ann. Eugen., Lond. *17*, 177, 1953.

Romanus, T.: Acta genet. *4*, 266–273, 1953

FREQUENCY OF CONSANGUINEOUS RELATIONS AMONG APPLICANTS FOR LEGAL ABORTION AND AMONG THEIR PARENTS IN SWEDEN

By TORSTEN ROMANUS, M. D. Upsala

More and more research in genetics is coming to rely on analyses of consanguineous relations. Their relatively high frequency among parents of recessive carriers – first demonstrated by *Lenz* [1919] – is one of the main reasons for this. Later *Dahlberg* [1947] went into the matter more deeply, and furnished more exact formulae. In the same paper he also stated that the size of isolates can be estimated if one knows the rate of cousin marriages. And in practice, as *Dahlberg* pointed out, only cousin marriages are any good, since statistics are not available for more distant degrees of consanguinity than first-cousin marriages – although they would have the same effects in principle – and the latter are merely available for some populations.

Fig. 1

Material.

In this investigation consanguineous relations of all kinds are considered. A new basis for an investigation of consanguinity is here tried, namely the frequency of consanguineous relations among women applying for legal abortion and among their parents. These women were all Swedish subjects who applied to the Royal Medical Board in the period from July 1, 1946, to June 30, 1950, for legal abortion on medical, eugenic or socio-medical grounds. The date July 1, 1946 was chosen because the law governing legal induction of abortion was then expanded to include socio-medical grounds as a valid indication. Information on consanguinity between the applicant and the expected child's father, and between the applicant's parents, are given in what are known as "Certificate A" and "Certificate C", the former being signed by a relative or relatives or an almoner and the latter by a qualified physician. Thus the veracity of any such statements is backed up by at least two independent persons. Relatives are probably in a position to really check up on the situation, whereas the physician merely can certify that the applicant has made such and such a statement. Women making more than one application in the period encompassed by this investigation were counted once only. The geographic classification was based on the place where the woman met the man. Those were rejected whose expected child would have been of doubtful paternity, or whose parents were unknown.

The data to be analysed statistically were extracted from the records by the author personally.

Representativeness of sample.

An enquiry conducted by the Royal Medical Board shows that the mean age of the applicants is 30 years and that the social stratification of those who apply for legal abortion and of their mates is about the same as for the population as a whole. Yet it is possible that the applicants selected included a preponderance of persons from the poorer sections of society and rather too few from the nobility where, of course, consanguineous marriages are probably rather frequent, which perhaps also results in too low a frequency of consanguineous relations in this investigation. On the other hand, however, the data were probably more complete than for average investigations of this type, because extramarital relations and illegal forms of intercourse, e.g. incest, were included. And, for purposes such as this, it seems

unlikely that known cases of consanguinity would deliberately be concealed. Concealment must otherwise be expected owing to the widely held opinion that consanguinity is something undesirable. This probably stems from the fact that intercourse with "near of kin" was forbidden by Mosaic law (Leviticius, Ch. 18, v. 6). In Sweden, too, first-cousin marriages were, up to 1845, allowed only by Royal dispensation; and in 1642 Axel Oxenstierna proposed an unconditional ban, to prevent Charles X Gustavus from marrying his cousin Queen Christina whom Oxenstierna wanted to marry his own son Eric! (*Almquist* [1937].)

A woman applying for legal abortion feels desperate and will not conceal a consanguineous relationship if she thinks it will strengthen her plea. This is amply confirmed by the records: sometimes very distant relations are mentioned (see table 3), even "relations" such as brothers-in-law.

Classification.

The series of women, comprising 15802 applicants, were classified according to their whereabouts, as follows: (see the map)

Group I: Stockholm with 703000 inhabitants (1.1.1948).
Group II: Gothenburg and Malmö, 337000 and 181000 inhabitants respectively, i.e. 419000 altogether.
Group III: Other towns except Visby.
Group IV: Gotland with Visby. Gotland, an island in the Baltic, must be regarded as a group of small isolates. Its area is about 3000 square kilometres and its population about 59000; the density of population therefore works out at about 19 inhabitants per square kilometre.
Group V: Rural parts of Jämtland, Västerbotten and Norrbotten Provinces. Group V must also be looked upon as a composite of comparatively small isolates, with about 2 inhabitants per square kilometre.
Group VI: Remaining rural Sweden.

The country as a whole has an area of about 410000 square kilometres, about 7000000 inhabitants, and a density of population of about 17 per square kilometre.

The applicants (= probands) have been thus classified in tables 1 and 2 and 3.

Results.

Table 2 shows that the frequency of consanguinity among the probands was 0.90 per cent for the country as a whole, and among their parents slightly higher, namely 1.13 per cent. The difference is

0.23 ± 0.11 per cent, and it proves almost wholly due to the difference in intercourse between first cousins, i.e. 0.69 – 0.41 corresponding to 0.28 ± 0.083 per cent, a statistically significant difference. The frequency of consanguinity was, as expected, lowest in Stockholm, 0.5 per cent, and highest for group V (a number of small isolates) with 4.6 per cent. For the latter group the frequency of first-cousin relationships was 1.5 per cent, or about 6 times that in Stockholm where it was 0.28 per cent. A comparison of the probands' "and their parents'" generations shows that first-cousin relationships were more common for the parental generation in all groups but group V and possibly group IV, i.e. the isolates. In the absence of information about the domicile of the parents when the

Table 1. The number of probands and of their parents with consanguineous relations.

Year		*Group I* Stockholm		*Group II* Gothenburg Malmö		*Group III* Other towns except Visby		*Group IV* Gotland with Visby		*Group V* Rural parts of Jämtland, Västerbotten and Norrbotten Provinces		*Group VI* Remaining rural Sweden		All Sweden				
		Total	with relationships	Total	with relationships	Total	with relationships	Total	with relationships	Total	with relationships	Total	with relationships	Total	First cousins	Second cousins	Other	Total with relationships
1946	Probands	319	2	221	2	276	0	3	0	15	2	229	6	1063	5	2	5	12
1/7-31/12	Parents		4		1		3		0		1		8		9	6	2	17
1947	Probands	749	2	519	2	775	7	28	1	74	6	683	14	2828	16	6	10	32
	Parents		6		10		9		1		3		17		25	12	9	46
1948	Probands	1263	10	479	6	1234	10	37	1	90	6	1027	11	4130	22	9	13	44
	Parents		11		5		14		0		2		13		32	6	7	45
1949	Probands	1509	7	501	2	1567	10	40	0	142	3	1201	15	4960	16	9	12	37
	Parents		8		5		15		0		2		16		29	6	11	46
1950	Probands	883	2	262	1	914	3	19	0	90	2	653	10	2821	6	6	6	18
1/1-30/6	Parents		9		2		5		0		1		7		14	3	7	24

Table 2. Frequency of the probands' and their parents' consanguineous relations.

Population	Number of probands	Relationships	Probands		Parents	
			Number	%	Number	%
Group I Stockholm	4723	first cousins	13	0.28	22	0.47
		second cousins	3	0.06	8	0.17
		other	7	0.15	8	0.17
		total	23	0.49	38	0.80
Group II Gothenburg and Malmö	1982	first cousins	10	0.50	15	0.76
		second cousins	1	(0.05)	4	0.20
		other	2	0.10	4	0.20
		total	13	0.66	23	1.16
Group III Other towns except Visby	4766	first cousins	17	0.36	27	0.57
		second cousins	4	0.08	8	0.17
		other	9	0.19	11	0.23
		total	30	0.63	46	0.97
Group IV Gotland with Visby	127	first cousins	1	(0.8)	0	—
		second cousins	0	—	1	(0.8)
		other	1	(0.8)	0	—
		total	2	(1.6)	1	(0.8)
Group V Rural parts of Jämtland, Västerbotten and Norrbotten Provinces	411	first cousins	6	1.5	2	0.5
		second cousins	3	0.73	2	0.5
		other	10	2.4	5	1.2
		total	19	4.6	9	2.2
Group VI Remaining rural Sweden	3793	first cousins	18	0.47	43	1.1
		second cousins	21	0.55	10	0.26
		other	17	0.45	8	0.21
		total	56	1.5	61	1.6
All Sweden	15802	first cousins	65	0.41	109	0.69
		second cousins	32	0.20	33	0.21
		other	46	0.29	36	0.23
		total	143	0.90	178	1.13

probands were born, we can only say that first-cousin relations were more frequent in the elder generation, in its turn a consequence of expanding isolates due to development of communications. In group V, an isolate, the frequency of consanguinity proves consistently higher in the probands' generation. An explanation could be that the probands were born in other parts of the country. It is interesting to note, however, that 84 of 90 women applying in 1950, i. e. 93 per cent, were born within the isolate. More likely, therefore, the

Table 3. Specification of relationships for those entered under "Other relationships".

6 probands.	Number
	4
hter and father	4
rnal uncle is the expected child's father	2
nd is maternal "half-aunt" of the expected child's father	1
nd and the expected child's father's father are cousins	2
nd and the expected child's father's mother are cousins	1
nd and the expected child's father are 'half-cousins'	1
nd's father and the expected child's father are cousins	1
nd's mother and the expected child's father are cousins	6
nd's father and the expected child's father are second cousins	1
nd's mother and the expected child's father are second cousins	2
nd's father and the expected child's father's father are second cousins	1
nd's father and the expected child's father's mother are second cousins	1
nd's mother and the expected child's father's father are 'half-sibs'	3
nd's mother and the expected child's father's father are second cousins	2
nd's maternal aunt and the expected child's father are second cousins	1
nd's father's mother and the expected child's father's father's mother are cousins	1
nd's maternal uncle and the expected child's father's mother's mother are cousins	1
nd's father's mother's mother and the expected child's father's mother's mother's father are sibs	1
nd's second cousin and the expected child's father are cousins	1
nd's father's father and the expected child's father's father's mother's mother are sibs	1
nd and the expected child's father are relatives of the fifth degree	1
nd and the expected child's father are 'distantly' related	6
hip given but not recorded	1

parents of 36 probands.	
band's father and proband's mother's father are identical	1
band's father and proband's mother's maternal uncle are identical	1
band's parents are third cousins	1
band's mother's mother and proband's father are cousins	5
band's mother and father are 'half-cousins'	1
band's father and mother's father are second cousins	1
band's mother and father's father are second cousins	1
band's mother and father's mother are second cousins	1
band's mother's mother and father's father are 'half-sibs'	1
band's mother's father and father's mother's father are sibs	1
band's mother's father and father's father are second cousins	2
band's mother and her father's mother's mother are cousins	1
band's mother is proband's father's cousin's daughter	2
band's father and mother are relatives of third or fourth degree	1
band's father and mother are relatives of the fifth degree	1
band's father's mother is proband's father's father's cousin's daughter	1
band's father and mother are 'distantly' related	13
ship given but not recorded	1

belief in the undesirability of consanguineous relations has persisted longer in the isolate. The applicants' generation has renounced this inhibition under pressure from geographic isolation.

It may be mentioned with regard to the frequency of first-cousin marriages in Sweden that *Dahlberg* has found a cousin marriage rate of 0.45 per cent for school children from central Sweden (Södermanland Province) and *Böök* [1948] 2.2 per cent in a northern Swedish isolate. Such figures agree rather well with that for the probands in this investigation.

The present-day decrease in consanguinity apparently depends on the decrease in first-cousin marriages while second-cousin marriages and other relationships remain the same as before and in rural parts possibly have become even more frequent. And, as *Dahlberg* pointed out, these kinds of consanguinity have in principle the same effect though it is more difficult to get reliable figures for them than for first-cousin marriages.

Summary.

The frequency of consanguineous relations between women applying for legal abortion in Sweden during the period 1946–1950 and their partners and between the applicants' parents was investigated. The material comprises 15 802 applicants with a mean age of 30 years. The frequency was found to be 0.90 per cent for the applicants and 1.13 per cent for their parents. The difference was due to a lower incidence of first cousin relations for the applicants and their partners (0.41 per cent) as compared to the parents of the applicants (0.69 per cent), while other types of consanguinity occurred with approximately the same frequency in both groups.

Résumé.

Les recherches concernent la fréquence des rapports consanguins observés en Suède parmi les femmes ayant fait une demande d'avortement entre les années 1946 et 1950 et également parmi leurs parents. Ces recherches comprennent 15 802 cas princeps âgés de 30 ans en moyenne. Les rapports consanguins chez les cas princeps s'élèvent à 0,90 % et chez leurs parents à 1,13 %. On observe une différence entre d'une part les rapports entre cousins germains s'élevant à 0,41 % pour les cas princeps et à 0,69 % pour leurs parents et d'autre part les rapports entre cousins issus de germains et autres membres de la parenté qui sont à peu près équivalents dans les 2 groupes.

Zusammenfassung.

Die Untersuchung berührt die Frequenz von Verwandtschaftsbeziehungen in Schweden zwischen Frauen mit Ansuchen um Unterbrechung der Schwangerschaft während der Jahre 1946–1950 und ihren Partnern, sowie unter den Eltern der Probandinnen und umfaßt 15 802 Probandinnen mit einem durchschnittlichen Alter von 30 Jahren. Zwischen den Probandinnen und ihren Partnern ergaben sich Verwandtschaftsbeziehungen mit 0,90 % und unter deren Eltern mit 1,13 %. Der Unter-

schied erstellt sich aus Beziehungen zwischen Vettern und Cousinen, welche unter den Probandinnen 0,41 % und unter deren Eltern 0,69 % betragen, während «anderweitige Verwandtschaft» in ungefähr gleichem Maße in den beiden Gruppen vorkommen.

REFERENCES

Almquist, J. E.: Acta Univ. lund. N.F., *33*, 1937. – *Böök, J. A.:* Hereditas *34*, 252, 1948. – *Dahlberg, G.:* Mathematical Methods for Population Genetics. S. Karger, Basel 1947. – *Lenz, F.:* Münch. med. Wschr. *66*, 1919.

Slater, E.: Acta genet. *4*, 273–280, 1953

SEX-LINKED RECESSIVES IN MENTAL ILLNESS?

By ELIOT SLATER

The present standing of genetics in psychiatry is now an assured one. For this we owe much to the contributions, both in subject matter and in methodology, of *Gunnar Dahlberg*, whose influence has been felt here as in much wider fields. The picture of today is greatly changed from that of the early years of the century, when research workers were largely fumbling in the dark, with no assured basis of factual knowledge, and but weak mathematical tools of analysis. This paper is the result of an examination of material collected by one of the early workers.

Soon after the turn of the century, *Frederick Mott*, the Director of the Pathological Laboratory of the London County Council Mental Health Service, began the collection of genetical material from the mental hospitals. In his "Heredity Index" a card-filing system was begun recording the names, diagnoses, dates of admission and discharge, etc., of all persons admitted to one or other of the L.C.C. Mental Hospitals who were known to have or to have had any other relative as a patient in one of these hospitals. It was on the basis of this material that *Mott* promulgated his now obsolete theory, the "Law of Anticipation". A grant from the Research Fund of the Institute of Psychiatry, Maudsley Hospital, has recently enabled a re-examination of some of the enormous amount of information which this Index contains; and a psychiatric social worker, Mrs. *Kate MacSorley*, was appointed to collect clinical records from the

files of the mental hospitals and to cooperate in their statistical analysis.

One of the groups of cases which we thought would repay examination was that of the avuncular relationships. The Index contained 233 pairs of persons so related, for which the clinical records could still be traced, distributed as follows (Table 1):

Paternal uncle – nephew	28
Maternal uncle – nephew	39
Paternal uncle – niece	34
Maternal uncle – niece	19
Paternal aunt – nephew	15
Maternal aunt – nephew	15
Paternal aunt – niece	33
Maternal aunt – niece	40

These figures are in themselves worthy of comment. There are 113 maternal relationships and 110 paternal relationships, when one would have expected the latter, owing to the greater probability of identical surnames, to have outnumbered the former. Aunts, however, with 103 representations, are outnumbered by uncles (120). Nieces (126) are better represented than nephews (97). These variations appear to be entirely haphazard, and convey no obvious meaning.

The next point which emerges from these figures is the very great preponderance of relationships between persons of the same sex. This is shown in Table 2:

	Uncles	Aunts	Total
Nephews	67	30	97
Nieces	53	73	126
Total	120	103	223

A χ^2-test yields a value of 16.09, which is very highly significant. Similar findings have been made before, e.g. by *Penrose*, and are capable of a variety of interpretations. Specific genetical factors which predispose to mental illness may, for instance, have different manifestation rates in the two sexes. In the present case, however, a genetical explanation is unsatisfying, as the sib relationships we have collected show no such excess of relationships between persons of the same sex. In this group we have brother-brother 58, brother-sister 130, sister-sister 65, the opposite-sexed relationships outnumbering the same-sexed relationships. It seems most probable that the informants who supplied the information to the mental hospitals, i.e. the relatives of the patients, were more likely to remember a second case of mental illness in the family if this relative were of the same sex as the patient being admitted; and that this tendency would have a more marked

effect if the forgetting of such cases were easy, as it would not be in the case of sibs.

Among the 8 different varieties of avuncular relationships there is one, the maternal uncle - nephew relationship, which is of especial interest genetically. For it is in this group that mental illness due to a sex-linked recessive gene would be most likely to appear. The importance of such genes in causing mental illness might have been disclosed by an excessive representation of this variety of relationship among all avuncular relationships; but in the material we have there is no sure sign of such an excess. The next test one might apply is the diagnostic one. Is there or is there not any particular variety of mental illness which is more common in the maternal uncle - nephew relationship than in the others?

The diagnostic classification of our material was a matter of considerable difficulty. The records were very old, sometimes dating from nearly a century back, and, judged by modern psychiatric standards, are very far from being sufficient. Nevertheless an attempt was made to diagnose the cases into the following groups:

Schizophrenic, paranoid (SP), catatonic (SC), hebephrenic (SH);
Affective (A), schizo-affective, manic-depressive, involutional;
Organic (O), arteriosclerotic, confusional, general paresis, paranoid, senile;
Epileptic (E);
Mentally defective (MD);

The numbers found in the various subdivisions of the affective and of the organic psychoses were so small that, in the tables which follow, these sub-classifications have been dropped.

Tables 3 and 4 show the various combinations of diagnoses between members of the older and of the younger generations:

		Maternal uncles							
		SP	SC	SH	A	O	E	MD	T
Nephews	SP	3	–	–	2	3	–	5	13
	SC	1	1	–	4	–	1	–	7
	SH	2	–	–	1	1	1	–	5
	A	–	–	–	–	1	–	–	1
	O	2	–	–	1	2	–	–	5
	E	–	–	–	–	1	–	–	1
	MD	3	–	–	–	1	–	3	7
	T	11	1	–	8	9	2	8	39

Table 3

		Uncles and aunts (other)							
		SP	SC	SH	A	O	E	MD	T
Nephews and Nieces	SP	16	8	3	3	8	1	–	39
	SC	6	14	3	8	16	1	–	48
	SH	5	2	2	8	5	2	1	25
	A	1	1	2	18	6	2	1	31
	O	3	–	1	6	8	2	–	20
	E	7	–	2	1	5	2	–	17
	MD	–	–	–	–	2	1	1	4
	T	38	25	13	44	50	11	3	184

Table 4

Looking first at Table 4, we see that there is a very marked preponderance of combinations of identical diagnoses. Schizophrenic psychoses go with schizophrenic psychoses, and moreover paranoid schizophrenias and catatonic schizophrenias each go with their like to a marked excess. Affective psychoses go with affective psychoses, and organic psychoses go with organic psychoses. The excess of such combinations of identities is statistically highly significant; but it is not proposed to go into the matter in this place any further. The tendency of like to go with like is no more than one would naturally expect on the general assumption that genetical factors play a part in the causation of mental illness.

Comparing Table 3 with Table 4, we see that there is only one feature in which Table 3 shows anything remarkable, and that is in the excess of cases of mental deficiency. We see a surprising number of cases in which mental deficiency has been paired with mental deficiency and also with paranoid schizophrenia. It seems quite possible that in these cases a sex-linked recessive gene was playing a causative role. This may be further examined by going to the clinical records, though unfortunately at this point the inadequacy of the reports becomes painfully apparent. Nevertheless short epitomes are appended.

Case 14. The maternal uncle was admitted to a mental hospital at the age of 36. He was said to have had heart trouble and emphysema. He was in good touch and clearly oriented, and spoke coherently. For the past 6 years he had a consuming passion for the Princess Royal, and had sent her two volumes of his own poems, and he believed his passion were reciprocated. When he learned that the Princess was ill, he went to Buckingham Palace, where he was apprehended by the police. On admission, he was mildly depressed but not emotional, and was decent, tractable, clean and tidy. He remained in hospital for about 5 years, occasionally protesting at his detention, before finally he could be discharged, now saying that he was too old for such passions. Throughout his stay there was no deterioration, and he was certainly not mentally defective.

The nephew was admitted to hospital at the age of 18, having marked his neck and his wrists with suicidal attempts. He was very tall, over 6 feet. He was found to be an emotionally unstable and impulsive feeble-minded individual, who grinned inanely, wandered restlessly about, was untidy and slovenly and much given to masturbation. In his short stay he seemed to deteriorate, but in 3 years was dead from pulmonary tuberculosis. Apart from being feeble-minded, it is possible that he had a superimposed hebephrenic schizophrenia.

Case 15. The maternal uncle was admitted to hospital at the age of 56 and again at 59. Before the earlier admission he had attempted suicide, but no record is available. All his life he was stated to have been mentally dull, had been given to alcoholism and had been in trouble with the police. On admission he laughed and

grimaced, and wandered aimlessly around. In talk with the doctors he was cheerful and garrulous, but had no understanding of the reason for his admission. Apart from feeblemindedness there was no gross psychiatric abnormality; and after a month he was discharged.

The nephew was admitted to hospital at 20. He had a history of having been in a school for truants, in Borstal and more than once in prison for various offences. On admission he was clearly oriented, but hallucinated; he heard men speaking of murdering him, of pulling the wall down. He was elated and said he could bring people back to life. He was well-behaved, but seemed to be feeble-minded; the hallucinations ceased to be observed, but he remained euphoric and insightless, untruthful, unreliable, homosexually inclined. He remained exactly the same person throughout his 7 years stay, before his parents arranged to take him home.

Case 21. The maternal uncle was admitted at the age of 27, his sister reporting he had been destructive and spiteful and had threatened to cut her throat. He spoke with his eyes closed, and burst into silly laughter. He soon settled down, but showed himself to be imbecile, unable to say how many days in a month or how many months in a year. In due course he was transferred unchanged, to an institution for mental defectives.

The nephew was also admitted to hospital at the age of 27. Two of his brothers had for some time been in hospitals for the defective, as they were imbeciles. The patient himself had shown no sign of mental defect, and had been regularly employed. Before admission he had started to court a girl, and had developed headaches and insomnia. On admission he was hostile and aggressive, said he had studied psychology, had cured himself of an inferiority complex and would now cure others. He was excitable and argumentative, readily assaulted others, and had odd mannerisms. Insulin coma treatment was begun but had to be broken off because of tachycardia. Nevertheless he steadily improved, and could be discharged recovered after 3 years.

Case 36. The maternal uncle was admitted to a mental hospital at the age of 28, supposedly suffering from "dementia", and six years later was transferred to an asylum for imbeciles. No other information is available about him.

The nephew was admitted to hospital at 42, in a state of restless excitement. All his life he had been weak-minded; and on admission is described as "degenerate" in appearance. His excitement was attributable to alcohol. After about a year he was discharged.

This man's sister was also in a mental hospital suffering from a catatonic schizophrenia, and was admitted there at the age of 20. Her mother said she had always been "queer", and lately worse, thinking that people were against her and that she was dead. On admission she was frantically excited, screaming at the top of her voice, at other times saying she had been mesmerised and turned into a dog, people controlled her thoughts and put strange ideas into her head, and she had seen devils in the sky accompanied by angels. She remained restless, or fixed in one attitude muttering to herself, at times depressed, at times impulsive. Her habits deteriorated, and she had to be tube fed. After some months she took a turn for the better, and was eventually discharged home after four years, though still retaining some strange religious ideas and some mannerisms.

Case 39. The maternal uncle was admitted to a mental hospital at the age of

57, and registered as "schizophrenia". He was still in the hospital 25 years later, but no other information is available.

The nephew was admitted at the age of 13. The history states that labour had been protracted and instruments used at his birth; he had attended a blind school. Father and mother were first cousins. One brother, an epileptic and an imbecile, had died at 17; a sister, imbecile and blind, had died at 18; another sister, imbecile and blind, had died at 14; a brother of 17 and a sister of 16 were alive and healthy. The patient was undersized, nearly stone blind, simple and weak-minded. He was restless, garrulous and noisy, actively hallucinated, seeing fires around him and fearing the ceiling would fall on him; he resisted examination, fearing he would be buried. His restlessness rapidly became worse, and before long he was running a high temperature. He died after a few months' stay of acute pulmonary tuberculosis.

Case 46. Here again, the uncle's paper are not available, having been destroyed (burnt with other records for fire-lighting, when the records office at the hospital was occupied by fire-watchers during the war); he is registered as having been an "epileptic imbecile", having had infantile paralysis at 4, but epilepsy in the first year of life. He was admitted to hospital at 16, transferred 8 years later to an institution for mental defectives. One of his brothers, therefore also a maternal uncle of the two nephews mentioned below, was admitted to the same mental hospital on the same day, then being aged 21, diagnosed as "imbecility with epilepsy", having been epileptic since the age of 3 months. He too was transferred with his brother to the same institution for defectives.

One nephew was admitted to hospital at the age of 32, having been healthy till then, and employed as a carpenter. For some time he had become increasingly strange, and before admission had tried to cut his throat. On admission, he was fully in touch, but had various delusions, that his sister had had intercourse with his brothers and his father, and was now pregnant. He talked incoherently, laughed nervously, seemed mildly depressed. After admission he became more vacant, at times confused, attacked attendants, at times said his mind was blank. His language became more strange. He said he had been admitted to hospital because of a "transparent table", which had been causing accidents. He suffered a progressive schizophrenic deterioration until his death of tuberculosis four years after admission.

A brother of the above was admitted to hospital at 25. The earlier part of the records is lost, but the latter part shows a deteriorated schizophrenic, self-absorbed, apathetic, talk so incoherent as to be almost unintelligible, at times hallucinated. Apart from imperceptible deterioration, there has been no change over the last 30 years, and he is still in hospital.

Case 58. The maternal uncle was admitted to hospital at the age of 26, a builder's labourer. He had recently developed delusions that there was a battery under his bed, and had become excited. He had made his way to Buckingham Palace, got past the guards, and only had been halted by facing rifles. Eight police constables struggled with him before he could be got away. In hospital he was unable to explain this episode, merely said he was a "stone-breaker". He was rambling, at times violent, refused food because it was poisoned. But he gradually improved, and could be discharged recovered after nearly two years.

The nephew was admitted to hospital at 16; he had done rather less than averagely well at school, but had since been employed in a factory, until he began to go deaf and to lose his sight. On admission he was dull and morose, liable to fits of

ungovernable temper lasting about an hour. There was complete bilateral nerve deafness, and bilateral corneal opacities due to interstitial keratitis. There is no record of a W.R. He remained in much the same state until his death from pulmonary tuberculosis two years later. In the final stages he is said to have been unable to articulate.

Case 61. The maternal uncle was admitted to hospital at 51, but the earlier part of the records has been lost; he remained in hospital until his death at 72. At 60 he is described as wandering in conversation, auditorily hallucinated, often excited and noisy, delusions of persecution by electricity. Apart from variable phases of excitement or apathy, he did not change, and never completely deteriorated in personality.

The nephew was admitted at 13, having been well till 6 months previously, though always excitable, irritable. He had become peculiar, had smashed crockery, threatened to stab his mother. On admission there was a lisp and a squint. He had phases of excitement in which he would bang his head on the floor. There was no subsequent change, and he did not deteriorate, but was always slovenly, incapable of the simplest mental operations, only employable at a very primitive level. He could finally be taken home.

Reviewing these clinical histories, it cannot be said that we have found anything very striking. The most unusual case is No. 39, in which a progressive dementing process, with onset in childhood, is found in four sibs, in three of them combined with blindness. Two of these four are female, and as in addition the parents were first cousins, we may be sure that the pathogenic gene involved was an autosomal recessive (unless it were partially sex-linked). In the other case in which physical abnormalities were involved, 58, the most probable diagnosis in the nephew is congenital general paresis. For the rest, we are left merely with the suggestion that sex-linked recessives may play a part in the causation of mental illness, of a schizophrenic kind, as well as mental deficiency; but that no very unusual clinical picture can, in that case, be expected.

Summary.

A genetical card index, instituted by *Mott* in the early years of this century, provided notifications of 233 pairs of persons, each of them admitted at some time to a London mental hospital, and related to one another as uncle (aunt) – nephew (niece). Pairs of like sex were in marked excess, but it is suggested that this was due to the mode of ascertainment. There was no obvious excess of maternal uncle–nephew relationships, but in that group there seemed to be an excess of cases of mental deficiency, either related to one another or to cases of paranoid schizophrenia. An examination of such clinical records as are available showed the existence of one family in which an early dementing process, sometimes combined with blindness, was probably due to an autosomal recessive gene, but otherwise revealed nothing very striking clinically. Although the suggestion arises that there may be

sex-linked recessive genes which can lead either to mental deficiency or schizophrenia, neither of the conditions so caused would be psychiatrically very distinctive.

Résumé.

Une collection de fiches génétiques, établie par *Mott* au début de ce siècle, a donné des renseignements sur 233 couples de personnes dont chacune avait, une fois ou l'autre, séjourné dans quelque maison d'aliénés à Londres et qui avaient les unes par rapport aux autres un certain degré de parenté p. e. oncle (tante) et neveu (nièce). Les couples du même sexe avaient une prépondérance marquée, mais ceci doit probablement être attribué à la manière dont on avait rassemblé les données. On n'observait pas de prépondérance marquée du lien de parenté entre oncle maternel et neveu, mais les cas d'imbécillité, apparentés soit les uns aux autres, soit aux cas de schizophrénie paranoïde, semblaient par contre être prépondérants dans ce groupe. Une analyse des fiches cliniques disponibles révélait l'existence d'une famille dans laquelle une psychose précoce, parfois aggravée de cécité, était due probablement à un gène autosomique récessif, mais à part ceci les fiches ne révélaient rien de très frappant du point de vue clinique. Bien que l'on puisse supposer qu'il existe des gènes récessifs liés au sexe, pouvant déterminer soit l'imbécillité, soit la schizophrénie, les états provoqués de cette manière ne doivent pas être très distincts les uns des autres du point de vue psychiatrique.

Zusammenfassung.

Ein genetisches Kartenregister, welches von *Mott* während der ersten Jahre dieses Jahrhunderts angelegt wurde, hat Aufschlüsse über 233 Paare von Individuen geliefert, von welchen jedes irgend einmal in einer Irrenanstalt in London aufgenommen worden war und welche einander gegenüber die Verwandtschaft Onkel (Tante)—Neffe (Nichte) besaßen. Paare gleichen Geschlechtes besaßen ein ausgeprägtes Übergewicht, welches aber wahrscheinlich der Art der Materialsammlung zuzuschreiben ist. Ein offenbares Übergewicht für Mutterbruder—Neffe-Verwandtschaft fand sich nicht vor; dagegen schien für miteinander verwandte für Schwachsinn oder in bezug auf Fälle von paranoider Schizophrenie ein Übergewicht vorhanden zu sein. Eine Analyse der zugänglichen klinischen Journale wies das Vorkommen einer Familie auf, bei welcher eine frühe Geisteskrankheit, zuweilen mit Blindheit verbunden, wahrscheinlich doch ein autosomales rezessives Gen verursacht war; jedoch enthüllten die Journale nichts, was klinisch gesehen besonders auffällig wäre. Obwohl es denkbar ist, daß es geschlechtsgebundene, rezessive Gene geben kann, welche entweder Schwachsinn oder Schizophrenie hervorrufen können, so dürften doch die Zustände, welche auf diese Weise verursacht worden sind, psychiatrisch gesehen nicht sehr deutlich von ähnlichen Bildern unterschieden sein.

REFERENCE

Penrose, L. S.: J. ment. Sci. *88*, 308, 1942.

Stern, C.: Acta genet. 4, 281–298, 1953

From the Department of Zoology, University of California
Berkely, California, USA

MODEL ESTIMATES OF THE FREQUENCY OF WHITE AND NEAR-WHITE SEGREGANTS IN THE AMERICAN NEGRO

By CURT STERN

The American Negro population represents a group which is unique genetically. Beginning with the latter part of the 17th century, for about two hundred years a variable and at times considerable amount of miscegenation took place between the North American inhabitants of African and European origin. During the last hundred years a much smaller but uninterrupted process of racial mixture has continued. Further interracial genetic heterogeneity has been introduced by miscegenation of Negroes and American Indians. This heterogeneity has been superimposed upon a conspicuous degree of genetic variability already present in those regions of Africa from which the majority of American Negroes were derived originally, as well as upon the heterogeneity of the Europeans.

As a result of these past and present processes the American Negro group undergoes continuously a complex segregation in respect to alleles of different racial origins. Among the segregants the White or near-White individuals have been the subject of particular discussion. Segregants whose more obvious phenotypic traits make them fall within the range of phenotypical variability of Caucasians, are potentially able to leave the relative sociological isolate of the American Negro community and to "pass as Whites". Estimates as to the number of segregants who have thus become members of the White community have varied from a few thousand to several millions. Most of these estimates reflect "more accurately the absence of adequate data than . . . the amount of passing occurring in the United States" (*Wirth* and *Goldhamer* [1944]). The only studies which are based on specific data are those of *Hart* [1921] and *Johnson* [1925];

see also *Day* [1932]. *Johnson* interpreted the fact that census data gave a striking deficiency of "Colored" males compared to females as evidence for a greater number of segregant males than females who have entered the records as "White". As interesting as this approach is, it has been pointed out that the census data, at least of the year 1910 used by *Johnson*, contain so many sources of error as to make them unsuited for an analysis of the problem (*Wirth* and *Goldhamer* [1944]). A similar situation exists for the census data as interpreted by *Hart*. He found, even after the necessary corrections for deaths, a striking deficiency of numbers of colored individuals in various age classes of the 1900 and 1910 censuses as compared to the corresponding ten year younger age classes of the 1890 and 1900 censuses. The combined apparent "loss" in all age classes during the decade from 1890 to 1900 constituted 301 000 individuals and in the decade from 1900 to 1910 355 000. No similar comparisons for more recent decades have been published. The reliability of these estimates is open to serious doubts intrinsic in the records but they still are the only ones based on a rational approach.

If the genetically relevant facts were known, independent estimates of the number of segregants could be made. These facts would have to include information on: (1) the frequency of alleles of Caucasian derivation in the American Negro, (2) the number, types, interactions and linkage relations of genes concerned with significant differences between Africans and Caucasians, as well as the genetics of relevant variation within each group, (3) genetic criteria for judging segregants as White or near-White, (4) the overall system of mating among American Negroes and the frequency and type of contemporary miscegenation. Listing these prerequisites makes our lack of knowledge apparent. While there is information on point (1), only fragmentary data are available on points (2) to (4). Nevertheless, an attempt will be made in the following pages to set up various models from which to derive consequences concerning the frequency of segregants.

The frequency of White alleles in the American Negro.

The proportion of genes of Caucasian derivation in the American Negro has recently been estimated by *Glass* and *Li* [1953]. On the basis of data on the frequency of several specific alleles of several loci, notably the R^0 allele, in American Whites, African Negroes and the American Negro these authors concluded that approximately 30 per

cent of the genes in the American Negro are of Caucasian derivation. This estimate leaves out of consideration the well established fact that a considerable number of American Indian genes have also been infused into the American Negro. The numerical value of this contribution is unknown. *Herskovits* [1930] reports that 27.2 per cent of a sample of American Negroes from Howard University stated they had partial Indian ancestry. *Meier* [1949] in an inquiry into the racial origins of Mississippi College Negroes obtained statements according to which 68.6 and 73.1 per cent of students born respectively in Mississippi or in nearby States had partial Indian ancestry. The difference in the information obtained by the two investigators has at least two reasons. First, there exists a true difference between the two groups studied. When African slaves arrived in larger numbers in the Eastern Seaboard States the number of Indians living there had already decreased greatly so that relatively limited miscegenation was possible. On the contrary, the Indian population of the Lower Mississippi Valley was considerable at the time of the introduction of Negroes into that region and miscegenation frequent. The second reason for *Meier*'s higher percentage is a purely secondary one. Between the times of *Herskovits'* and *Meier's* study nearly a generation had elapsed. Since the offspring of any parents, one of whom had Indian and the other no Indian ancestry, would state his ancestry as containing an Indian element, the number of individuals with such ancestry must obviously increase with each generation.

Glass and *Li*'s calculations regarding the Rh^o allele can not throw light on the problem of Indian admixture since the frequencies of this allele are similar in American Whites and Indians. Thus their estimate of 30 per cent non-African admixture in the American Negro is compatible with any proportion of White to Indian ancestry. Data on the frequencies of the I^o allele, which is responsible for the 0 blood group do not support the assumption of a high degree of American Indian admixture. According to the compilation of *Glass* and *Li* the frequencies of the I^o allele in African Negroes and North American Whites are very similar, values varying between 0.66 to 0.72 for the former, and being about 0.67 for the latter. For American Negroes frequencies of between 0.66 to 0.69 have been determined. Since the "unmixed" American Indians had an I^o frequency of nearly 1.0 any appreciable contribution by them to the American Negro should have increased his I^o frequency above those of the Africans and Caucasians. It would be interesting to compare in more detail allele frequencies

of other genes in the African, Caucasian, American Indian and American Negro. It might then be possible to estimate with greater accuracy the relative contributions of the three groups.

In the absence of definite knowledge as to the overall genic contribution of the American Indian to the American Negro population the fractions $p=0.7$ and $q=0.3$ for the allele frequencies of African and non-African elements will be used in the subsequent treatment. Moreover, the non-African element will for simplicity's sake be designated as "White". In addition, some estimates of the results will be presented assuming that the fraction of non-African genes in the American Negro is as low as 20 per cent.

The inheritance of skin color.

The most obvious difference between Negroes and Whites is in pigmentation of the skin. Within limits, a light skin is therefore a minimum criterion for a White or near-White segregant.

Genetic estimates of the frequencies of such segregants depend on the model set up to account for the inheritance of skin color in Negro-White crosses. Such a model was first proposed by *Davenport* [1913]. He found that (a) he could match any skin color of Whites, Negroes, and hybrids in Jamaica, Bermuda, and Louisiana by an appropriate mixture of black, red, yellow and white segments blended together by the rotation of a color top; (b) a continuous range of skin color exists from very light to very dark; (c) five peaks of frequencies of pigmentation types in terms of the relative area of the black disc segment seemed discernible in his population measurements; and (d) the offspring of $F_1 \times F_1$ matings contained a few white and black individuals in addition to a variety of intermediate types. On the basis of these findings *Davenport* postulated a 2 gene pair scheme for the difference between the colors of Whites and Africans, assigning to the Whites four genes with a low base contribution to pigmentation and to the Africans four genes with identical, additive action toward increased pigmentation. Although this model seemed reasonably satisfactory for the interpretation of matings between various color types of the hybrid populations many discrepancies had to be overlooked.

Later *Todd* and *van Gorer* [1921] proposed as a measure of skin pigmentation the sum of the relative area of the black disc segment plus 59 per cent of that of the red segment, since the red disc actually contains 59 per cent black. *Todd* and *van Gorer*'s method has been used extensively and forms the basis of most later evaluations of skin

color. *Barnes* [1929] has transformed *Davenport*'s data to *Todd*'s scale. This transformation results in the disappearance of four of *Davenport*'s five frequency peaks which in any case were hardly significant features of his populations. *Barnes* analyzed the family data of *Herskovits* and collaborators [1930, and unpubl.] on skin pigmentation in the American Negro as well as *Davenport*'s older report. Among her findings were the following: (a) for the middle range of average pigmentation of the parents the average pigmentation of the children was like that of the parents; (b) on the average, light parents have somewhat darker children and dark parents somewhat lighter children than themselves; (c) the overall regression of average pigmentation of offspring on average pigmentation of parents was of the order of 2.43 per cent on 3 per cent groups; (d) only slight evidence for a minor degree of dominance in skin color inheritance was apparent; (e) no simple genetic model could be fitted to the data as a whole.

Recently, *Gates* [1949] suggested a three-pair gene scheme to account for the inheritance of skin colors in Negro-White crosses without however attempting to draw quantitative conclusions from his model. How little we know is emphasized by a passing statement of *Penrose* [1953] that "the number of genes involved might be anything between 2 and 20."

The following discussion is based upon setting up a series of alternative models, comparing their consequences with certain empirical data, and showing that only a limited number of the models lead to expectations which are in approximate agreement with the facts. This does not exclude that other types of models can be designed which would fit the observations at least equally well.

The observed distributions are shown in figure 1 for two samples the contents of which partly overlap. While basically similar there are significant differences between them. *Barnes*' curve is based on 2391 individuals measured by various observers. All data have been corrected for age changes in pigmentation by treating the population as if it represented individuals of the age group 34–39 only. This population contains 365 individuals especially selected for higher economic and professional status, a group which was found to be significantly lighter than the average American Negro. *Herskovits*' curve is based on 2710 individuals, not corrected for age and without a selected subgroup. This latter feature makes it more representative. The lack of age correction tends to give this curve a somewhat higher pigmentation average than that of the corrected one. Besides, the presence of

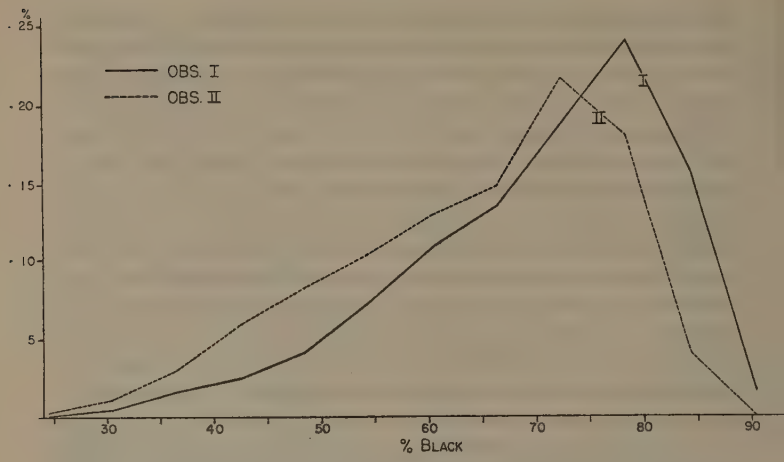


Fig. 1. Frequency distributions for skin pigmentation in the American Negro as measured by total corrected per cent black in a matching color top. Obs. I: *Herskovits* [1930]; obs. II: *Barnes* [1929].

the selected lighter subgroup probably accounts for a considerable fraction of the surplus of lighter types, in the range from 25 to 60 per cent black, in *Barnes*' curve. Therefore, *Herskovits*' distribution will be used in the following comparisons with expected distribution.

The expected distributions are based upon various models all of which share 3 assumptions: (1) All genes involved in increasing pigmentation beyond a light basic figure have equal, additive effects as expressed in terms of per cent of black in a matching color top. (2) The scale of per cent of black is linearily related to the number of pigment increasing genes present in the different genotypes. (3) Random mating and genetic equilibrium. On the background of these assumptions models of various pairs of genes, n, will be investigated, namely for n equal to 2, 3, 4, 5, 6, 7, 10 and 20 pairs. Each of these models gives a limited number of classes of pigmentation, from 5 for the 2 pair model to 41 for the 20 pair model. The models of equal additive multiple gene pairs fit the fact of agreement between average pigmentation of parents and offspring within the middle range. They do not take account of the observed, though minor regression at the extremes of the range.

In order to account for the continuous nature of the actual varia-

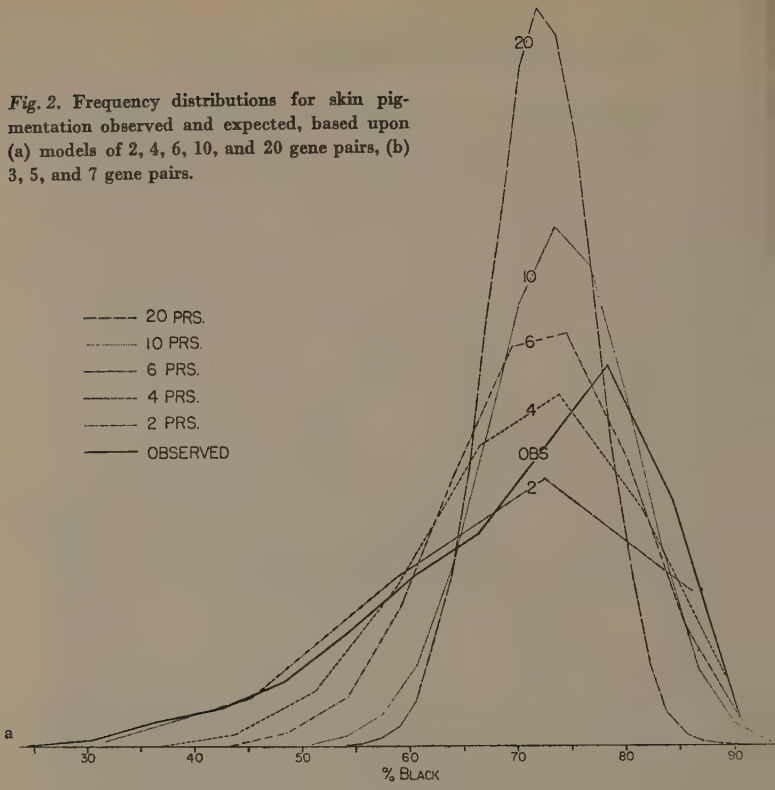
tion of pigmentation it is necessary to invoke environmental influences as well as additional genetic modifiers of undetermined kind and number. These modifiers must be thought of as also being responsible for part of the variation in skin color within "unmixed" Caucasians and Africans. In part the major, additive genes responsible for the interracial difference may also play a role in the intraracial variance, particularly in case of models with higher numbers of such genes. Thus, in the 5 pair model, all Caucasians may not be completely free from those alleles which cumulatively are responsible for the dark pigmentation of the African, but some may contain 1, 2 or even 3 of these alleles. Similarly, Africans may contain in lower frequency some of the alleles which are prevalent among Caucasians.

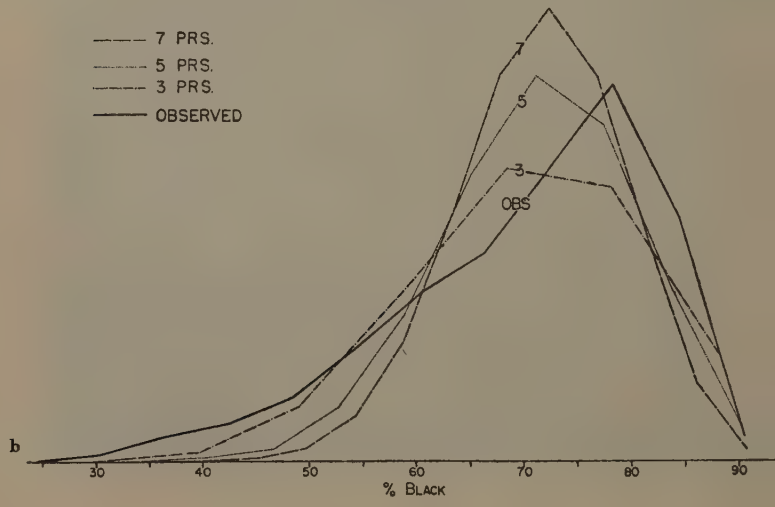
The expected distributions were obtained from expansions of the binomial $(p+q)^{2n}$ where $p = 0.7$, $q = 0.3$, and n has values varying from 2 to 7, and is also taken as 10 and 20. The total empirical range of pigmentation extends over 68 units, from 25 to 93 per cent. This range was divided into as many equal subranges as there are numbers of terms in each expanded binomial. The value of each term was then entered on the midpoint of its corresponding range. In doing this, each value was adjusted to the number of classes of its binomial so that the total area under the various distributions remained constant.

A comparison between the observed and the expected distributions (figures 2a, b) shows two kinds of deviations common to most though in various degrees. In the lighter ranges, up to nearly 60 per cent black, and in the darkest range, above 85 per cent, the frequencies of observed types usually exceed those of the expected. In addition to these general deviations there are striking differences in the middle range between some of the expected and the observed distributions. The observed frequencies greatly exceed the distributions for 2 and 3 pairs of genes but fall far short of those for 20 and 10 pairs. The models for 4, 5, 6 and—least—7 pairs approach, to varying degrees the empirical findings.

The deviations between observations and expectations have two different kinds of causes, namely imperfections in sampling and intrinsic disagreement between the various models and the actual circumstances. Some of the excess of observed light as well as dark types is due to partial positive assortative mating in respect to pigmentation (*Herskovits* [1926]). Some may be due to the presence of minor isolates, in an interbreeding sense, of groups of the American Negro who are most, or least, infused with Caucasian genes. Among

Fig. 2. Frequency distributions for skin pigmentation observed and expected, based upon (a) models of 2, 4, 6, 10, and 20 gene pairs, (b) 3, 5, and 7 gene pairs.





the intrinsic differences between the models and reality there may be a systematic one, namely the assumption of lack of dominance over the whole range of pigmentation. *Barnes'* finding of regression of the pigmentation average of the offspring on those of the parents, primarily at the two extremes, could be interpreted by assuming some dominance of genes acting toward light and dark pigmentation respectively. Finally, a non-linear scaling relation between per cent black in the color top reading and the additiveness of genic actions is likely to be involved.

In spite of the deviations the relative agreement between the observed values and those expected on the basis of the 4, 5 and 6 gene pair models is rather suggestive. The fit would presumably be even better if observations were available from the contemporary population of the American Negro. A whole generation has passed since the time of *Herskovits'* sample. It is likely that by now a closer approach to genetic equilibrium has been attained.

Estimates of segregation frequencies.

In the preceding section it has been shown that models employing 4, 5 or 6 pairs of genes give an approximate fit to the observed distribution of pigmentation types. The consequences of these models for the problem of White or near-White segregation will now be determined. It will be useful to add to this determination a consideration of the less likely models employing 2, 3, 7 or 10 pairs of genes so that a wider range can be explored. Only the 20 pair model will be exluded from further treatment since it is so excessively out of line with the facts.

For the purpose of this discussion a White or near-White segregant has to be defined in quantitative terms. Restricting the analysis first to skin color, two partly independent criteria are available, namely measured percentage of black and frequency of extreme pigmentation segregants in the F_2 progenies.

The criterion of measured percentage of black is based on the relation between percentage of black in segregants and in the unmixed White population. *Blackwood* [1930] has measured the pigmentation of 331 Whites from various parts of the United States primarily including individuals of British and Scandinavian ancestry. While no detailed data have been published the mean and standard deviation for the percentage of black were given as 32.54 ± 4.29. A smaller sample of 50 White cadavers from the city of Cleveland gave 39.8 ± 10.01

(*Todd*, *Blackwood* and *Beecher* [1928]). It may be surmised that the Cleveland sample included an appreciable number of relatively dark persons, of Mediterranean and South-Eastern European ancestry. A comparison of the means and standard deviations of the two samples of Whites suggests that the Cleveland group has a skewed distribution with an excess of extreme dark deviants. If, somewhat arbitrarily, the American White population is assumed to be 35 ± 7, then more than 97 per cent of the Whites would possess pigmentation in the range below 49 per cent black and only 0.62 per cent of the population be darker than 52.5. It may therefore be regarded as unlikely that many individuals of the American Negro population with pigmentation in excess of about 50 per cent black will be considered as White or near-White segregants. In other words, such segregants in general will have to be lighter than 50 per cent black.

It was the frequency of extreme pigmentation segregants in F_2 progenies which served as the basis for *Davenport*'s 2-pair gene model. Altogether no more than 5 sibships were measured. Three of the 32 F_2 individuals fell in the color range of Whites and one in the range of Africans. Since a 2-pair hypothesis would require one sixteenth of each for White and Negro segregants, the agreement with the observed frequencies seemed good. In reality the five sibships were clearly not homogeneous. Moreover the pigmentation values for the three White segregants were at the dark extreme of the range for Whites as defined by *Davenport*. In the following discussion therefore, a considerable range of segregation percentages, namely from 3 to 25 per cent, for White or near-White color segregants in F_2 will be accepted as compatible with *Davenport*'s limited observations. Since there is a positive correlation between frequency of expected segregants in F_2 and the range of pigmentation, from very light toward dark, which is included in any specific definition of segregants, a mutual check on the validity of the two criteria for segregants is provided. It will be employed below.

For F_2 progenies, the frequencies of the various combinations of dark pigmentation genes, from 0 for the lightest to 2n for the darkest combination, are given by the expansion of the binomial $(0.5+0.5)^{2n}$. These frequencies have been computed for the various values of n to be considered. Since the frequency of White or near-White segregants was chosen so as to lie within the range of 3 to 25 per cent of the F_2's, for each gene model the frequencies of as many of the gene combinations at the light end of the pigmentation range were added as to

provide a sum of frequencies between 3 and 25 per cent. In the 2 pair model, for example, the consecutive percentage frequencies of F_2 combinations containing 0, 1, 2 etc. genes for dark pigmentation are 6.25, 25.0, 37.5 etc. Here the 0 class with 6.25 per cent lies within the defined range 3–25 per cent, while the sum of the frequencies of the first two groups, 31.25 per cent lies outside this range. Similarly, for the 5 pair model, the percentage frequencies for the combinations with 0, 1, 2, 3, 4 etc. genes for darkness are 0.1, 1.06, 4.4, 11.72, 20.51, etc. respectively. This means that both the 0 class and the combined 0 and 1 classes would account for less than 3 per cent segregation in F_2, that combination of all classes from 0 to 4 would give more than 25 per cent segregation (namely 37.7 per cent), while the sums of combinations 0–2 and 0–3 with 5.47 and 17.19 per cent segregation respectively, would both fall within the defined range. These and all those other sums of combinations for the various models which fit in the defined range are entered in table 1, column 3. It is seen that they lie between 3.5 and 21.2 per cent. The frequencies of those light genotypes or their sums which lie outside the range of 3 and 25 per cent are either as low or lower than 2.9 or as high or higher than 29.2 per cent (columns 4, 5).

In column 6 of table 1 the upper limits of pigmentation percentages are entered as they correspond to the various sums of genotypes.

Table 1. Estimates of percentage frequencies of White and near-White pigmentation segregants in a random mating population of the American Negro according to different models and specifications.

gene pairs	dark genes in segregants O to x	segregants in F_2	segregants in F_2 when segregants O to (x—1)	segregants in F_2 when segregants O to (x+1)	upper limit % black	segregants in population if q = 0.3	segregants in population if q = 0.2
2	0	6.2	—	31.2	38.6	0.81	0.160
3	0–1	10.9	1.6	34.4	44.4	1.02	0.160
4	0–1	3.5	0.0	—	40.1	0.13	0.008
4	0–2	14.4	—	36.3	47.7	1.13	0.123
5	0–2	5.5	1.1	—	43.5	0.15	0.008
5	0–3	17.2	—	37.7	49.7	1.05	0.086
6	0–3	7.3	1.9	—	45.9	0.17	0.006
6	0–4	19.4	—	38.8	51.1	0.95	0.058
7	0–4	9.0	2.9	—	47.6	0.16	0.005
7	0–5	21.2	—	39.5	52.2	0.82	0.038
10	0–7	5.8	2.1	—	51.0	0.13	0.002
10	0–8	13.2	—	29.2	54.2	0.51	0.010

It is seen that the 2, 3, 4 and 5 pair models imply upper limits for White and near-White color segregants below 50 per cent black, compatible with the criterion developed earlier. Among the sums of combinations relating to the 6 and 7 gene pair model only one for each remains below the 50 per cent limit and both of the sums for the 10 pair model surpass it. Had percentages of F_2 segregation above 25 per cent been permitted, the range of the pigmentation would have gone beyond the limit of 50 per cent in each case.

Within the models chosen, the genotypes which result in White or near-White pigmentation segregants have now been established. In order to obtain values for the frequencies of the sum of these genotypes not in an F_2 group but in the American Negro population when at genetic equilibrium, these sums were computed for the relevant terms of the binomial $(0.7+0.3)^{2n}$. They are shown in the next to the last column of table 1. The frequencies vary between 0.13 and 1.13 per cent. Should *Glass* and *Li*'s estimate of 0.3 for the White allele frequency in the American Negro be too large, the estimates of the frequencies of White segregants would be disproportionally too large. This is shown in the last column in which a White allele frequency of 0.2 has been assumed. Consequently, the frequencies of segregants vary from as low as 0.002 to no more than 0.16 per cent. If the White allele frequency were 0.25 no segregant frequency would have reached 0.5 per cent.

Other traits than pigmentation which enter into the phenotypic designation of an individual of the American Negro population as a White segregant remain to be considered. The most important ones of these are width of nose and thickness of lips, although others such as yellowness of sclera, or color and shape of hair also play a role. No data suitable for a genetic interpretation are available but a preliminary estimate of their significance can be made from the known means and variances of at least two of the traits in American Whites and Negroes (*Todd* and *Lindala* [1928]). For 100 Whites, width of nose was found to be 34.9 mm $\pm$ 3.68 (σ) as compared to 100 Negroes with 42.4 $\pm$ 3.68. For thickness of lips the values were 11.5 $\pm$ 3.70 (Whites) and 21.1 $\pm$ 4.92 (Negroes). It follows that one half of all Negroes have wider noses than 98 per cent of the Whites, and one half of all Negroes thicker lips than 99.5 per cent of the Whites. The ability of a segregant to pass over into the White population would obviously depend upon the phenotypic combination of different traits according to a differential scale.

Medium expression of Negroid facial features might play only a minor role if present in a very lightly pigmented individual, while even submedian expression of Negroid facial features might restrict the potential passing ability of a somewhat darker segregant. Since the majority of White pigmentation segregants belong to the darker range of such segregants the last consideration has to be given considerable weight. As an overall guess, subject to possible empirical modification, it will be assumed that a maximum of only 25 per cent of all White pigmentation segregants will have a combination of other phenotypes such as facial features, hair properties, etc., which will enable them to be classified as White. Accordingly, in order to obtain estimates of the frequencies of these individuals the values of the last two columns of table 1 would have to be multiplied by 0.25.

Assortative Mating.

The estimates obtained depend, among others, on the assumption of random mating. *Herskovits* [1926] has shown that this assumption is close to reality for nose width and thickness of lips since for these traits there is only a negligible amount of correlation between spouses. In respect to pigmentation, however, significant deviations from random mating were observed. There was a tendency toward positive assortative mating and, in addition, it was found that more often the wife was the lighter one of a couple than the husband. (This latter observation implies a surplus of unmarried darker women, but no observations on this point have been assembled). Using data from *Herskovits* and his collaborators, *Barnes* also found nonrandomness between spouses in regard to pigmentation (table 2, column 4). A comparison with column 2 shows that there was positive

Table 2. Assortative Mating according to degree of pigmentation in the American Negro. Based on data from 326 marriages tabulated by *Barnes* [1929] (table 16). "Light" < 45 per cent black; "medium" = 45–67 per cent; "dark" > 67 per cent.

Marriages	Random expectation	Absolute positive assorted expectation	Observation
light × light	3.8	35.0	6
light × medium	28.7	—	38
light × dark	33.8	—	20
medium × medium . .	54.7	133.5	70
medium × dark	128.7	—	89
dark × dark	76.1	157.5	103

assortative mating between the three color ranges distinguished. However, comparison with column 3 indicates how low the degree of assortative mating still is as compared with completeness.

Assortative mating increases the frequencies of White segregants over those calculated for random mating. It is believed to be unlikely that at the present time this increase would as much as double these frequencies but no thorough analysis of this problem has been undertaken.

Some of the models and subsidiary assumptions explored in this study have given a moderate fit to empirical data on the distribution of pigmentation in the American Negro. The 4–6 gene pair models have led to a maximum expectation of 1.13 per cent White pigmentation segregants and, including considerations of assortative mating, of a maximum of $2 \times 0.25 \times 1.13 = 0.565$ per cent of total phenotypic segregants. These figures are valid for the first generation of a random mating population at genetic equilibrium. With an American Negro population of the size of the present day, namely 15 million, a maximum of about 85000 individuals per generation of more than 25 years, or a yearly quota of less than 3400 potential total segregants would be involved. Many of the estimates presented in table 1 would correspond to very much lower numbers. Moreover whatever the genetically derived estimates are, it is believed by students of the American Negro that only a small fraction of total White segregants actually desire to pass out of the Negro community.

On the other hand, it should be emphasized that the American Negro population most likely has not yet reached genetic equilibrium. That this was true for the last generation is suggested by the following facts. Except in the case of multiple effects of a gene, no correlation between unrelated phenotypes would usually be expected within a population which is at equilibrium. There was, however, a slight positive correlation in *Herskovits'* sample between width of nose and thickness of lips. (No measurement of correlation between degree of pigmentation and facial features was attempted.) Furthermore, the frequency of light segregants among the group studied by *Herskovits* was considerably higher than predicted by all except the 2 pair model of color inheritance used in the estimates presented.

It may be assumed that a greater approach to randomness has been made since *Herskovits'* studies. It would, therefore, be of great interest to know what the correlation between the various features

is in the present generation of the American Negro, and how the frequency of light segregants compares at present with the model estimates. Such information would enable us to judge to what degree the over-simplified model estimates can be regarded as approaching reality.

Pigmentation frequencies after panmixis in the total population.

The models suggested for the genetic determination of pigmentation differences between Caucasians and Africans can be used to calculate expected distribution frequencies for the array of pigmentation types if after complete panmixis within the whole American population an equilibrium would be obtained. The results of such calculations are presented in table 3. They are based on the models of 3, 5, and 7 gene pairs and, alternatively, on frequencies of 0.3 and 0.2 of White alleles in the American Negro. The present ratio of 9 : 1 for Whites to Negroes in the American population has been assumed as constant yielding allele frequencies for dark genes in the total population of q' = 0.07 and 0.08 respectively. The table shows

Table 3. Frequencies of all types of pigmentation segregants in a random mating American population according to different models for gene pairs and of dark allele frequencies (q'). (For each model of gene pairs upper row q' = 0.08, lower row q' = 0.07).

Gene pairs	Number of dark genes										
	0	1	2	3	4	5	6	7	8	9	10 and more
	$\times 10^{-3}$	$\times 10^{-3}$	$\times 10^{-1}$								
3 pairs	60.6	31.6	6.9	8×10^{-3}	5×10^{-4}	2×10^{-5}	3×10^{-7}	—	—	—	—
3 pairs	64.6	29.2	5.5	6×10^{-3}	3×10^{-4}	1×10^{-5}	1×10^{-7}	—	—	—	—
5 pairs	43.4	37.8	14.8	3.4×10^{-2}	5×10^{-3}	5×10^{-4}	4×10^{-5}	2×10^{-6}	6×10^{-8}	1×10^{-9}	1×10^{-11}
5 pairs	48.4	36.4	12.3	2.5×10^{-2}	3×10^{-3}	3×10^{-4}	2×10^{-5}	8×10^{-7}	2×10^{-8}	4×10^{-9}	3×10^{-12}
7 pairs	31.1	37.9	21.4	7.4×10^{-2}	1.8×10^{-2}	3×10^{-3}	4×10^{-4}	4×10^{-5}	3×10^{-6}	2×10^{-7}	$<10^{-8}$
7 pairs	36.3	38.1	18.7	5.6×10^{-2}	1.2×10^{-2}	2×10^{-3}	2×10^{-4}	2×10^{-5}	1×10^{-6}	6×10^{-8}	$<10^{-8}$

that in each model case the great majority of the panmictic population falls within the range of variation of White pigmentation. Near-White segregants are rare and dark individuals almost absent.

I wish to end this paper with a plea for new data. For many years very little has been added to our knowledge of the genetics of interracial differences as far as they concern anthropologically important external traits. There are regions, as for instance certain parts of Brazil, where relatively easily ascertainable first generation, back cross, and second generation hybrids between Africans and Caucasians are frequent (*Pierson*, [1942]). Here seems to exist an important, but neglected field of study.

Postscript.

After the manuscript had gone to press *R. R. Gates* has elaborated on his model [1949] of three pairs of genes for pigmentation differences of Negroes and Whites (Studies of Interracial Crossing, II: A new theory of skin color inheritance. Intern. Anthrop. and Linguist. Review 1:15-67, 1953). *Gates* proposes, on the basis of "many observations" of individual pedigrees "some of which are detailed in the present contribution" that the three pairs of genes "are weighted for pigmentation, as follows: R = 6, S = 2, T = 1," where RRSSTT = 18 produces the darkest and rrsstt = 0 the lightest color.

When *Gates'* model is used to compute the distribution frequencies of color types in the American Negro striking deviations from the observed distribution become apparent. There is (1) an even larger deficiency of expected lighter types than was found to exist for the best fitting models discussed above, (2) a very great surplus of expected individuals of the darkest group, and (3) a bimodality of expectation in which the second darkest class is only slightly more than half as frequent as the next lighter class and somewhat less than one third as frequent as the darkest class. These deviations make it unlikely that the new model with its assumed considerable inequality of genic actions is close to reality, If, as seems a priori likely, some inequality of action exists of the different pairs of genes concerned with pigmentation differences, it is probably only of minor extent. The population-genetic analysis of *Gates'* model illustrates the need for simultaneous use of pedigree and population data.

Summary.

Distribution frequencies of color types in the American Negro are compared with expectations derived from various models of additive gene pairs involved in the inheritance of pigmentation differences between Caucasians and Africans. Models of 7, 10, and 20 gene pairs give an increasingly poorer fit to the data. Models of 2 and 3 gene pairs also give great deviations from observation. Models of 4, 5 and 6 gene pairs agree best with observations but with considerable deviations.

Using criteria of permissible frequencies of White and near-White color segregants in F_2 progenies and of pigmentation type of such segregants, estimates are presented for the frequencies of the segregants in the Negro population.

Additional considerations lead to assumptions regarding the fraction of the

pigmentation segregants who would approach Caucasian phenotypes in other but skin color characteristics.

Calculations of distribution frequencies of pigmentation in the total American population after panmixis suggest a virtual disappearance of the Negro.

Résumé.

Les distributions de fréquence des différentes types de pigmentation chez les nègres américains ont été comparées aux distributions attendues, calculées théoriquement en prenant comme point de départ des paires de gènes à effet semblable pouvant déterminer la transmission des différences de pigmentation entre les caucasiens et les africains. Des calculs basés sur 7, 10 et 20 paires de gènes s'accordent de moins en moins avec les observations. On observe aussi de grands écarts dans les observations si l'on utilise 2 ou 3 paires de gènes. On obtient la meilleure concordance si l'on base les calculs sur 4, 5 et 6 paires de gènes, mais les écarts sont encore considérables.

Utilisant des criteria de fréquences permissibles des ségrégants blancs et presque blancs dans les progénies F_2 et des genres de pigmentation de ces ségrégants, l'auteur fait des estimations du taux de telle ségrégation dans la population nègre. Des considérations supplémentaires conduisent aux suppositions concernant la fréquence des individus qui pourraient se rapprocher du type caucasique par d'autres caractères que par la couleur de la peau. Des calculs des fréquences de pigmentation dans toute la population américaine indiquent en cas de panmixie une disparition presque complète du type nègre.

Zusammenfassung.

Die Frequenzverteilungen in Bezug auf verschiedene Hautfarbentypen amerikanischer Neger sind mit erwarteten Verteilungen verglichen worden, welche, auf Grund verschiedener Genmodelle bezüglich der Vererbung von Pigmentierungsdifferenzen zwischen Kaukasiern und Afrikanern berechnet wurden. Berechnungen, welche sich auf 7, 10 und 20 Genpaare gründen, stimmen zunehmend schlechter mit den Beobachtungen überein. Ebenso liegen große Abweichungen vor, wenn man zwei oder drei Genpaare zu Grunde legt. Die besten Übereinstimmungen erhält man mit 4, 5 und 6 Genpaaren. Jedoch verbleiben beträchtliche Abweichungen. Unter Benutzung gewisser Kriterien bezüglich der Häufigkeiten weißer oder nahezu weißer Segreganten in F_2 Familien wird die Häufigkeit solcher Segreganten in der Negerpopulation geschätzt. Weitere Erwägungen leiten zu Annahmen bezüglich des Bruchteils der Pigment-Segreganten, die sich kaukasischen Phänotypen in anderen Eigenschaften annähern. Berechnungen von Frequenzverteilungen der Pigmentierung in der gesamten amerikanischen Population bei Panmixie deuten auf ein praktisch völliges Verschwinden des Negertypus hin.

REFERENCES

Barnes, I.: Hum. Biol. *1*, 321–381, 1929. – *Blackwood, B.:* J. roy. anthrop. Inst. *60*, 137–168, 1930. – *Davenport, C. B.:* Heredity of Skin Color in Negro-White Crosses. Carnegie Inst. Wash. Pbl. *188*, 1913. – *Day, C. B.:* Harv. Afr. Stud. *10*, 1–127, 1932. – *Gates, R. R.:* Pedigrees of Negro families, Blakiston, Philadelphia 1949. – *Glass, B.* and *C. C. Li:* Amer. J. Hum. Genet. *5*, 1–20, 1953. – *Hart, H.:* Univ. Ia. Stud. Child Welf. *1*, 1–137, 1921. – *Herskovits, M. J.:* Proc. nat. Acad. Sci. Wash. *12*,

587–593, 1926; Columbia Univ. Contr. Anthrop. *11*, 1–283, 1930. – *Johnson, C. S.:* Opportunity *3*, 291, 1925. – *Meyer, A.:* Amer. J. Physic. Anthrop. Ser. 2, *7*, 227–239, 1949. – *Penrose, L. S.:* Review of *Dunn, L. C.* and *Th. Dobzhansky*. Ann. Eug. *17*, 252–253, 1953. – *Pierson, D.:* Negroes in Brazil, 1–392, Univ. of Chicago Press, 1942. – *Todd, T. W., B. Blackwood* and *H. Beecher:* Amer. J. Physic. Anthrop. *11*, 187–204, 1928. – *Todd, T. W.* and *L. Van Gorer:* Amer. J. Physic. Anthrop. *4*, 239–260, 1921. – *Todd, T. W.* and *A. Lindala:* Amer. J. Physic. Anthrop. *12*, 35–121, 1928. – *Wirth, L.* and *H. Goldhamer:* The Hybrid and the Problem of Miscegenation in Characteristics of the American Negro (O. Klineberg, edit.), p. 253–369, Harper and Brothers, New York 1944.

Waardenburg, P. J.: Acta genet. *4*, 298–312, 1953

Aus der anthropogenetischen Abteilung des Niederländischen Instituts für Präventive Medizin in Leiden

ZUM KAPITEL DES AUSSEROKULAREN ERBLICHEN NYSTAGMUS

Von Dr. P. J. WAARDENBURG (Chef der Abteilung)

Es hat seit einigen Jahren den Anschein, als ob das Problem des erblichen Nystagmus endgültig gelöst sei, und zwar in dem Sinne, daß diese Anomalie ohne Ausnahme rezessiv- oder mehr dominant-X-chromosomal übertragen werde. Es ist wahr, daß man die Mehrzahl der veröffentlichten Stammbäume in dieses Schema hineinpassen kann. Man darf daraus jedoch nicht folgern, daß dies nun einwandfrei für alle Fälle erwiesen sei.

Es dürfte nützlich sein, das Gesamtmaterial in dieser Hinsicht noch einmal zu überprüfen. Dazu habe ich alle jene Fälle entfernt, die für Albinismus oculi mit sekundärem Nystagmus verdächtig sind. Dürfen wir deshalb aber annehmen, daß der primäre Nystagmus sich nur X-chromosomal vererbt?

Kitahara [1929] und *Lenz* [1932] haben das vermutet, jedoch nicht bewiesen. Aber trotzdem habe ich mich im Jahre 1935 für die überwiegende Mehrzahl der Fälle auf ihre Seite gestellt. Noch immer

aber ist die Frage nicht vollkommen gelöst. *Kitahara* hat seine Vermutung damals auf 5 eigene Sippen und auf 50 Sippschaften des Schrifttums gegründet. Seine eigenen Fälle waren dazu teilweise (z. B. Fam. 4 und 5) ganz ungenügend. Er fand die wichtige Tatsache, daß das Merkmal fast nie von Vater auf Sohn vererbt wird – im japanischen Nystagmusmaterial findet sich dieser Fall nur einmal bei *Kurata* [1923]. *Lenz* hat – klinisch wohl unhaltbar – Polyallelie mit drei verschiedenen phänotypischen Unterschieden der Allele angenommen, wovon zwei sich dem normalen Gen gegenüber rezessiv und eines, das die schwerere Anomalie verursache, dominant verhalte. Zur Erklärung der Ausnahmefälle, von denen *Lenz* eine geringere Anzahl annahm als es realiter im Schrifttum gibt, hat er vermutet, daß sie entweder unzuverlässig gewesen seien, oder daß es sich um Verwandtenehen und Übertragung mütterlicherseits gehandelt habe.

Gerade die nicht unbeträchtliche Häufigkeit der regelwidrigen Fälle hatte mich noch im Jahre 1932 dazu veranlaßt, vorläufig drei Vererbungsmodi beim extraokularen Nystagmus anzunehmen: a) eine autosomale Dominanz, b) eine X-chromosomale Dominanz mit derart häufiger, unregelmäßiger Penetranz, daß Rezessivität vorgetäuscht werden kann, c) eine autosomale Rezessivität auf Grund seltener Fälle von familiärem Nystagmus ohne Augenabweichungen in blutsverwandten Ehen und in Ehen von normalen, nicht verwandten Eltern. In beschränktem Sinne habe ich mich *Nettleship* [1909, 1911] angeschlossen, der eine unregelmäßig-dominante, autosomale «ambisexual» und eine rezessiv-X-chromosomale Form annahm.

Nach wie vor bin ich bei dieser Auffassung geblieben, habe aber annehmen müssen, daß die autosomal-dominante Form weitaus die seltenste ist. Ich verfüge nur über eine eigene Beobachtung bei Vater und Sohn, die vielleicht auch auf andere Weise zu erklären wäre, und finde des weiteren eine Mitteilung von *Fattovich* [1936] über eine Sippe mit kleinem, sehr raschem Nystagmus bei 11 von 43 Mitgliedern durch 4 Generationen, der direkt und regelmäßig von Vater oder Mutter auf Söhne und Töchter übertragen wurde. Dabei überwiegt die Vererbung von Vater auf Sohn so sehr, daß ein X-chromosomaler Erbgang ausgeschlossen ist.

Was den autosomal-rezessiven Erbgang anbelangt, verfüge ich augenblicklich über drei persönliche Fälle. Die Literaturfälle kann ich kaum beurteilen, da sie meistens zu lückenhaft sind, um sonstige Erbgänge auszuschließen.

Es bleibt jetzt übrig, wenn möglich die Ausnahmefälle des

Schrifttums richtigzustellen. *Kitahara* und *Lenz* haben diesen Versuch nicht unternommen. Ich bemühe mich schon seit Jahren darum, so weit es möglich ist, die nötige Korrektur anzubringen, und ich glaube, daß mir das größtenteils gelungen ist. Die Regelwidrigkeit, die richtiggestellt werden soll, besteht in der Feststellung, daß Väter und Söhne beide an Nystagmus leiden, oder – noch unbegreiflicher – daß normale Väter die Anomalie auf Söhne oder Töchter übertragen haben.

Am leichtesten ist es für mich gewesen, aus dem eigenen Material und demjenigen von *Hemmes* einige Irrtümer zu entfernen, zumal, da ich auch noch imstande gewesen bin, mehrere Stammbäume von *Hemmes* durch neu hinzugekommene Fälle zu ergänzen. Ein Punkt, in dem *Hemmes* und ich uns schon gleich von *Nettleship* haben trennen müssen, bestand darin, daß wir bei beiden von ihm angenommenen Erbgängen gelegentliches Kopfwackeln feststellen konnten, während *Nettleship* meinte, daß das nur bei der ambisexuellen, «autosomalen» Form vorkäme. Zweitens konnte ich die Vermutung von *Hemmes* entkräften, daß es sich bei den Zweigen LL und MM, wobei der letztere auf Grund anamnestischer Mitteilungen als autosomal angesehen wurde, um zwei verschiedene Erbgänge in einer Sippschaft handle. Ich konnte nämlich in dem sogenannten rezessiv-X-chromosomalen Zweig LL zu dem einen betroffenen Mädchen, das *Hemmes* gefunden, einige hinzufügen, wodurch es ausgeschlossen ist, daß hier ein autosomaler Erbgang durch Translokation in einen X-chromosomalen Erbgang – oder umgekehrt – stattgefunden hat. Auch der anamnestisch vermittelte Zweig MM paßt zum X-chromosomalen Erbgang.

Nun ist zu gleicher Zeit noch ein weiterer, wichtiger Unterschied mit *Nettleship* darin gegeben, daß der X-chromosomale Erbgang des außerokularen Nystagmus viel eher ein dominanter mit unregelmäßiger Penetranz als ein rezessiver ist, und daß letzterer nicht einmal für den Nystagmus des Augenalbinismus gilt, da er gewöhnlich oder immer intermediär verläuft. Der rezessive X-chromosomale Erbgang ist ein scheinbarer, der dadurch vorgetäuscht werden kann, daß viele Stammbäume so klein sind.

Versuch einer Korrektur der Regelwidrigkeiten.

1. Stammbaum *Waardenburg*, bei *Hemmes* als DD angegeben: der anamnestisch als betroffen angegebene Sohn eines befallenen Vaters hat sich bei Nachprüfung als normal ergeben.

2. *Hemmes* (Stammbaum H. H.). Es ergab sich, daß der links im Stammbaum als normal eingezeichnete Groß- und Urgroßvater von 4 befallenen Enkeln und Urenkeln, wobei zweimal durch nicht befallene Frauen weitervererbt wurde, nur auf anamnestische Daten beruhte, wobei der Nystagmus wahrscheinlich übersehen worden war (Abb. 1).

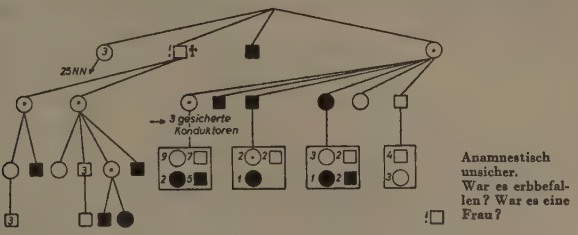


Fig. 1. Hemmes [1924] Stb. HH.

3. *Hemmes* (Stammbaum J. J.). Links unten ist ein befallener Mann durch einen normalen Vater und Großvater mit seinem befallenen Urgroßvater verbunden. Es ergab sich, daß der Nachkomme keinen Nystagmus bei Blick vorwärts hatte, sondern nur einigermaßen in Endstellungen, was ihm selber in keiner Weise bewußt war. Wahrscheinlich lag hier eine andere Anomalie vor (Abb. 2).

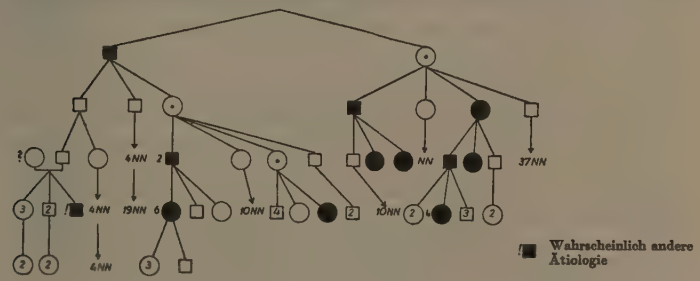


Fig. 2. Hemmes [1924] Stb. I. I.

4. *Lans* ([1908] Stammbaum M bei *Hemmes*). *Hemmes* hat irrtümlicherweise die Stammeltern verwechselt. Da wurde also nicht von Vater auf Sohn – wie *Hemmes* angibt – sondern von Mutter auf Sohn vererbt. In den zwei letzten Generationen konnte ich aber Angaben von *Lans* richtigstellen. Ein anscheinend befallener Sohn eines befallenen Vaters war frei von Nystagmus. Dagegen zeigte ein Mädchen in der letzten Generation Nystagmus, während *Lans* es als nicht befallen verzeichnete. Dieser letzte Fehler ist aber ohne prinzipielle Bedeutung. Ich konnte den Stammbaum mit 3 neuen, behafteten Knaben ergänzen, von denen der älteste schon wieder einen normalen Sohn und ein befallenes Mädchen hat, wobei nichts Regelwidriges vorlag. Soweit die Beobachtungen, die meine eigene Umgebung anbelangen. Ich komme jetzt zu den weiteren Streitigkeiten in der Literatur.

5. *Wood* [1892] (Stammbaum B bei *Hemmes*). H. hat den Stammbaum so gezeichnet, daß eine befallene Frau die Anomalie mittels eines normalen Sohns auf eine Enkelin überträgt. Es ergibt sich aber, daß der befallene Stammvater 10 normale Söhne haben soll und eine befallene Tochter, und daß eine Enkelin – wobei nicht angegeben wird, ob sie von dieser Frau oder von einem ihrer normalen Brüder abstammt – wieder betroffen ist (wahrscheinlich das erste, und dann würde alles stimmen) (Abb. 3).

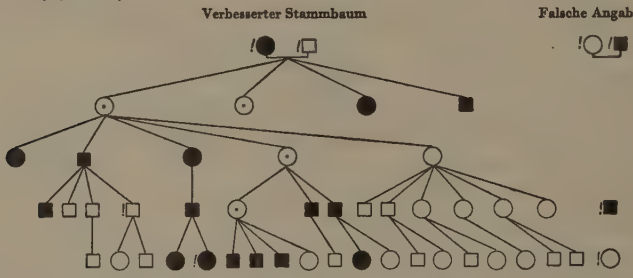


Fig. 3. Lans [1908], *Waardenburg* [1924].

6. *Radloff* [1909] (Stammbaum P bei *Hemmes*). In den letzten Generationen hat ein befallener Vater einen befallenen Sohn. Hier ergibt sich folgendes: Der Vater hatte einen langsam pendelnden Nystagmus *mit Kopfwackeln*. Von seinen 5 Kindern ist die älteste Tochter normal, die zweite ist ergriffen wie der Vater, die dritte Tochter zeigte sehr schwache nystagmusartige Bewegungen ohne Kopfwackeln. Dann kommt ein Sohn mit einem etwas deutlicheren, aber *noch sehr geringgradigen* Nystagmus, ebenfalls *ohne Kopfwackeln*. Ein jüngerer Sohn und eine jüngere Tochter sind normal (Abb. 4). Hier kommen vier Erklärungsmöglichkeiten in Betracht:

a) da es sich bei den leicht befallenen Geschwistern um ein anderes klinisches Bild handelt, könnte es auch genetisch eine andere Ursache haben,
b) der Vater hätte eine Verwandte geheiratet,
c) es würde eine durch non-disjunction X-chromosomlose Eizelle von einer X-chromosomhaltigen Samenzelle befruchtet, so daß ein befallener Sohn entstand.
d) es könnte eine Translokation von einem X-Chromosom an eine Autosom stattgefunden haben.

Es wäre sehr wichtig zu wissen, ob dieser Sohn das Leiden weiter vererbt hat und ob die sub b) geäußerte Vermutung wahr ist.

7. *Eugenie von Kibort* [1910] hat in einer neurologischen Studie zwei Stammbäume K. und R. in der Landschaft Davos ausgearbeitet, die bis ins 18. Jahrhundert zurück verfolgt werden konnten und wobei ursprünglich keine Verwandtschaft zwischen beiden Sippen bekannt war. Der Autorin gelang es jedoch eine Verbindung zwischen beiden Sippen aufzufinden dadurch, daß eine im Jahre 1798 geborene Elisabeth K., deren Mutter eine geborene R. war, ihren Namen K. auf einen außerehelichen Sohn übertrug, der seitdem der Stammvater des Zweiges K. wurde. Im Zweige R. kam bei *Kibort* nichts Regelwidriges vor. *Semadeni* hat diese Sippen 1939

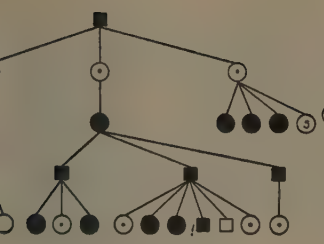


Fig. 4. Radloff [1909].

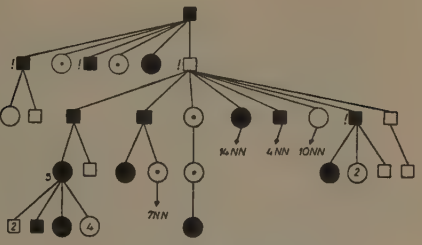


Fig. 5. E. v. Kibort [1910].

weitergeführt. Dabei hat sich nun leider ein Ausnahmefall mit Vererbung von Vater (F_6 9) auf Sohn im Zweig R eingeschlichen. Im Zweige K. kamen vom Anfang an mehrere Unstimmigkeiten vor, hauptsächlich in den älteren anamnestisch vermittelten Generationen (Abb. 5). Seit der Ergänzung *Semadenis* sind immer noch einige Male Angaben von Vererbung von Vater auf Sohn, einmal durch 3 Generationen hindurch, vorhanden. Es ist eine große Befriedigung, daß *Hanhart* die Riesenarbeit geleistet hat, diese wertvollen, sehr ausgedehnten Stammbaumzweige von Fehlern zu reinigen, so daß einwandfrei erwiesen werden konnte, daß auch diese Sippen dem unregelmäßig-dominanten X-chromosomalen Erbgang genau folgen. Einmal hatten zwei befallene Eltern vier befallene Töchter ohne weitere Komplikation.

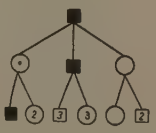


Fig. 6. Kurata [1923].

8. *Kurata* [1923] (nach *Komai*). Der Nystagmus ist hier bei 3 Männern in 3 Generationen vorhanden, einmal bei Vater und Sohn (Abb. 6). Es ist nicht bekannt, ob der Vater eine Blutsverwandte geheiratet hat, wie das in Japan viel vorkommt. Es wird angegeben, daß der befallene Großvater, durch eine phaenotypisch normale Tochter 3 befallene Enkel und 2 normale Enkelinnen erhält, daß eine andere äußerlich normale Tochter 3 normale Kinder bekommt (2 Söhne und 1 Tochter) und daß ein Sohn des Großvaters ebenfalls Nystagmus zeigt. Dieser Mann hat 3 Söhne und 3 Töchter ohne Nystagmus. Man sieht sich genötigt, außer Konsanguinität der Eltern in diesem einzigen Ausnahmefall entweder anzunehmen, daß eine Translokation stattgefunden hat, oder daß Non-disjunction bei der Großmutter vorgelegen hat, oder aber, daß der Nystagmus des Vaters oder des Sohnes zu einem anderen klinischen Typus gehört. Ein Fehler bei der Wiedergabe oder ein Druckfehler wäre schließlich auch noch möglich.

9. Stammbaum *Wherry* (*Nettleship* [1911]). Dieser Stammbaum ist am schwierigsten mit X-chromosomaler Vererbung in Einklang zu bringen (Abb. 7). Die Stammutter I_2 soll nach Aussage von III_{13} normal gewesen sein. Nach der hiesigen Auffassung wäre sie dann eine Konduktorin gewesen. Ihr ältester Sohn II_1

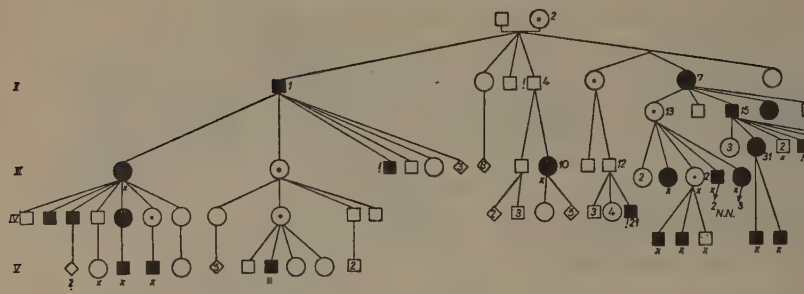


Fig. 7. Wherry [1911].

starb 85 Jahre alt anno 1905 und hatte sowohl Nystagmus als Kopfwackeln. Das wird wohl ganz gesichert sein, da er wenigstens 8 Nachkommen mit Nystagmus hatte und da seine Schwester II_7 (Partner eines weiblichen Drillings) selber befallen war und wenigstens 12 Nachkommen mit Nystagmus hatte.

Die erste Unstimmigkeit im Stammbaum beruht auf der Angabe, daß II_1 einen wieder befallenen Sohn III_3 hat. Ich zweifle daran, ob das zuverlässig ist. Anamnestisch wird mitgeteilt, daß dieser Sohn kinderlos verheiratet in Amerika lebt und daß seine Augen «move slightly». Da in dieser Sippe mehrere Male die nicht selbst beobachteten Fälle als positiv angegeben sind, wenn die zurategezogenen Ärzte, die die Patienten sahen, über «well marked» oder «very marked» Nystagmus berichteten, und da im allgemeinen in dieser Sippe der Nystagmus ein lebendiger war, möchte ich annehmen, daß dieser Sohn, bei dem nicht von Kopfwackeln wie bei seinem Vater geredet wird, nicht befallen war, oder daß seine Augenunruhe eine andere Ursache hatte. Sonst stimmt alles, was über die Nachkommen von II_1 berichtet wird sehr gut mit den Erwartungen überein.

Als zweiter Widerspruch ergibt sich, daß II_1 einen normalen Bruder II_4 haben soll, der wieder eine befallene Tochter gezeugt hat. Auch das ist nicht überraschend. Über diesen Bruder jedoch hört man sonst nichts als «Mark normal». Es mag deshalb bei dieser ganz ungenügenden Berichterstattung ein leichter Nystagmus übersehen oder vergessen worden sein, zumal, da der Mann nicht augenärztlich untersucht wurde. Es mag aber auch sein, daß die Tochter, von deren 5 oder 6 Söhnen keiner befallen war, selber eine andere klinische Form aufwies. *Nettleship* sagt nämlich, daß sie ausgezeichnet sah und nur «a slight nystagmus» besaß.

Die dritte noch größere Regelwidrigkeit ist, daß III_{12}, der ebenfalls in Amerika wohnte, auf einer Englandreise von Verwandten darüber befragt, angab, daß – obwohl er selber als normal im Stammbaum figuriert – unter 8 Kindern (4 Töchtern und 4 Söhnen) ein Sohn (IV_{21}) mit Nystagmus vorkommt («said by the father to have the moving eye»). Wenn man jedoch bedenkt, daß von einem 16jährigen Mädchen (IV_{17}) nach Aussagen einer Kusine bewegliche Augen gemeldet werden, während *Nettleship* im Gegenteil nur nervöses periodisches Lidkneifen feststellte, darf man m. E. den Nystagmus bei dem genannten bezweifeln, um so mehr, als seine 7 Geschwister davon frei sind.

Die vierte und letzte Unstimmigkeit in dieser Sippe beruht auf der Angabe, daß ein befallener Sohn (III_{15}) der obengenannten sicher befallenen Frau II_7 unter 5 Töchtern zwei und angeblich unter 4 Söhnen ebenfalls zwei befallene aufweist. Das letztere ist nicht möglich und ließe sich auch kaum durch zweimalige Nondisjunction erklären. Vom Vater ist nur bekannt, daß er 45 Jahre alt 1901 starb, und daß die Augen Bewegungen aufwiesen («eyes moved»). Er hatte dunkle Komplexion, Myopie und Sehbeschwerden. Wie werden nun diese Söhne beschrieben? Vom älteren Sohn (IV_{34}) der nicht ärztlich beobachtet wurde, sagen die Mutter und eine befallene Schwester (IV_{31}), die selber wieder 2 befallene Söhne hat, daß seine Augen wie diejenigen der Schwester (nicht also wie diejenigen des Vaters) beweglich sind. Liest man genau nach, was über diese sichere Konduktorin gesagt wird, dann findet man: «slight lateral nystagmus *varying* under observation and *sometimes absent*, sight seems good but not recorded». Man darf deshalb höchstens annehmen, daß genannter Bruder (IV_{34}) eine ganz leichte Form der Anomalie, wenn überhaupt, aufweist, die anderer Herkunft sein kann. Wenn nicht, so ließe sich in diesem Fall non-disjunction vermuten. Der jüngere Sohn (IV_{36}) wurde im Alter von 16 Jahren von *Nettleship* gesehen. Von ihm gilt folgendes: es gibt ein «very fine rapid nystagmus when first looked at, but *it passes off under observation*». Auch dieser Fall ist somit ein unsicherer, man darf nicht als verbürgt annehmen, daß hier die erbliche Sippenform des Nystagmus vorliegt. Die zweite befallene Schwester (IV_{35}) hatte den Nystagmus nur fakultativ, was bei einer Heterozygotin immerhin möglich ist, aber für einen hemizygotisch befallenen Mann etwas sonderbar wäre. Es bleibt eine Schwierigkeit in dieser Familie, daß von den beiden befallenen Söhnchen von IV_{31} der älteste «slight nystagmus at times when not looking attentively» hatte und der jüngste mit «a bad constant lateral nystagmus of rather wide range» beschrieben wird. So sind nicht alle Fälle in dieser Sippe klinisch oder phänotypisch gleich und deswegen sind für Nicht-Beobachter die vorgelegten Fragen nicht einwandfrei lösbar.

Wenn wir jedoch bedenken, daß es *Hanhart* möglich gewesen ist, alle Strittigkeiten, die noch im Stammbaum der großen Davoser Familie von *Kibort* nach der näheren Bearbeitung von *Semadeni* übriggeblieben sind, zu lösen, so daß diese Familie jetzt ebenfalls ganz genau in das unregelmäßig-dominante X-chromosomale Familie jetzt ebenfalls genau in das unregelmäßig-dominante X-chromosomale Schema hineinpaßt, dann kommt es mir am wahrscheinlichsten vor, daß diese angloamerikanische Sippe auch hierher gehört und wir können nur bedauern, daß sie noch nicht wieder erneut untersucht und ergänzt worden ist. Ich nehme vorläufig an, daß die nichtstimmenden Fälle III_3, III_{10}, IV_{31}, IV_{34}, IV_{36} und vielleicht auch noch der schon stimmende Fall V_{21}, unzuverlässig sind, und daß sie höchstwahrscheinlich irrtümlicherweise als mit der uns beschäftigenden erblichen Form des Nystagmus behaftet angegeben sind. Wenn das nicht der Fall sein würde, bleibt nur noch übrig, daß es neben der X-chromosomalen Form, noch eine nicht-rezessive autosomale Form gibt, die sich regelmäßig dominant vererbt wie im Falle *Fattovich* [1936] und vielleicht bei *Kurata*.

In der neueren Literatur finde ich nur noch diese Unstimmigkeiten:

10. *Aguilar* [1937]. In einer Sippe mit 5 durch 4 Generationen befallenen Männern und 8 befallenen Frauen besitzen 4 dieser Männer zusammen 10 Söhne, davon 9 normal und einer wieder befallen. Es ist möglich, daß bei den vielen anamnestischen Angaben dieser Fall fälschlich dasteht, oder daß eine der schon oben genannten Erklärungsmöglichkeiten gilt.

11. *Käser* [1941]. Dieser Autor hat gelegentlich Nystagmus in Endstellungen bei Söhnen befallener Väter gefunden, also etwas, was mit dem obengenannten Fall *Hemmes* (J. J.) vergleichbar ist und m. E. gegen gleiche Erbveranlagung spricht. Der Autor betrachtet auch diese leichte Anomalie, die er bei 3 männlichen (zweimal Söhne befallener Väter) und bei 3 weiblichen Personen (einmal Tochter eines befallenen Vaters) antraf als Manifestation des gleichen familiären Leidens, wobei dann jedoch X-chromosomaler Erbgang nicht mehr aufrechtzuerhalten wäre. Es kommt außerdem noch ein Knabe mit Nystagmus (V_3) vor, dessen Mutter und Großvater mütterlicherseits, der befallene Brüder hat, als normal angegeben sind. Offensichtlich sind in dieser Sippe andere Nystagmustypen hineingeschlichen.

12. *Allen* [1942] fand 11 Personen (5 Männer und 6 Frauen) durch 4 Generationen befallen, dabei einmal Großvater, Sohn und Enkel! Hierfür könnten alle schon vorher genannten Erklärungsmöglichkeiten geltend gemacht werden. Keine ist jedoch erwiesen.

Zusammenfassend läßt sich auf Grund umfassender Nachprüfungen jetzt behaupten, daß es mir und *Hanhart* in den meisten Sippen gelang, falsche Angaben, die häufig in den ältesten anamnestisch vermittelten Generationen vorkamen, richtigzustellen. Nicht alle Strittigkeiten konnten überall einwandfrei entfernt werden, diejenigen bei *Kurata*, *Wherry*, *Aguilar*, *Allen* nur vermutungsweise.

Es ist ebenfalls merkwürdig, daß das neuere Schrifttum uns außer der Mitteilung von *Fattovich*, von *Aguilar* und von *Allen* keine weiteren Befunde gebracht hat, die nicht in Einklang mit dominant-X-chromosomaler Vererbung wären. Die von *Knighton* [1929], *Cox* [1936], *Pellman Glover* [1937], *Barigozzi* [1939], *Leroy Billings* [1942], *Mac Gregor* [1946], *Wilbur Rucker* [1949, in einem Zweig mit Koppelung an partieller Farbenblindheit], *Cuendet* und *Della Porta* [1949] und in Kombination mit *Franceschetti* [1950] veröffentlichten Fälle stimmen sehr gut mit der Erwartung überein.

So ergibt sich, daß beim extraokularen erblichen Nystagmus überwiegend X-chromosomaler Erbgang vorliegt.

Es ist fünfmal vorgekommen, daß ein Mann mit Nystagmus zwei Ehen geschlossen hat. In drei Fällen wurde der Nystagmus in beiden Linien weitervererbt (*Cox* [1916], *Waardenburg* [1924], *Wilbur Rucker* [1949], in zwei nur in einer derselben (*Wilbur Rucker* [1946, 1949]. Das letzte stellt keine Unregelmäßigkeit dar, aber beruht wahrscheinlich darauf, daß die Anzahl der Töchter oder der verheirateten Töchter des befallenen Mannes in der einen Ehe nicht groß genug ist (zwei Töchter, davon eine verheiratet, zusammen nur eine Enkelin bei *W. Rucker* [1946], beim selben Autor 1949 in der einen Ehe nur ein Sohn). Derartiges gilt ebenfalls für die von *Rucker* angegebenen männlichen befallenen Zwillinge [1946]. Nur einer hat

die Anomalie weitervererbt, der andere hatte von seinen beiden Töchtern nur 3 Enkel.

Cuendet und *Della Porta* publizierten einen von *Franceschetti* eingehend besprochenen Stammbaum, wo zwei befallene weibliche, nicht näher angegebene Zwillinge das Merkmal beide auf Nachkommen vererbt haben.

Wenn wir mit der neuen Auffassung der X-chromosomalen Vererbung Recht haben, folgt daraus:
1. daß der Nystagmus *niemals* von Vater auf Sohn vererbt wird,
2. daß die Nystagmusanlage vom Vater auf *alle* Töchter übergeht,
3. daß die Anlage von Konduktorinnen auf die Hälfte der Söhne und der Töchter vererbt wird. Je regelmäßiger die Dominanz, oder anders gesagt, je höher die Penetranz ist, um so deutlicher wird zutagetreten, daß wirklich die meisten Töchter eines befallenen Mannes von neuem befallen sind. Bei vollkommener Penetranz müßten alle Töchter selber befallen sein, bei weniger vollkommener Penetranz müßten sich die verheirateten Töchter, wenn ihre eigene Kinderzahl nicht all zu klein ist, wenigstens als Überträgerinnen erweisen.

1. Zum *ersten* Punkt sei bemerkt, daß wir diesen nach obengenannten Berichtigungen als befriedigend erreicht betrachten dürfen. Sowohl für die gelbe (Japan) als für die weiße Rasse gilt, daß das praktisch überall stimmt. Wir besitzen schon große Zahlen (in Japan haben 6 behaftete Männer zusammen 13 Söhne, die alle normal sind). Bei der weißen Rasse finde ich 67 behaftete Männer, die zusammen 154 Söhne haben, welche alle normal sind. Von 149 Töchtern sind 54 wieder betroffen und 41 sicher Konduktor.

2. Wenn ich alle von *Hemmes* zusammengefügten Literaturdaten bei der weißen Rasse in dieser Hinsicht nachprüfe, so ergibt es sich, daß in 19 Sippen 37 befallene Väter 88 Töchter haben. Von diesen Töchtern ist erwiesen, daß 54 als Konduktor auftreten, dadurch, daß sie, obwohl phänotypisch-normal, wieder befallene Nachkommen erhielten. Ihnen gegenüber stehen 34 unverheiratete, nichtbefallene Töchter, die folglich keinen Aufschluß in dieser Frage geben können und eine verheiratete Tochter (Stammbaum Y, *Nodop*), die zwei normale Söhne und drei normale Töchter erhielt, was jedoch nicht beweist, daß sie selber nicht belastet war. Im japanischen Material (*Mikamo*, *Toyada*, *Kitahara* Fam. 1., 2. und 3.) haben 5 an Nystagmus leidende Väter zusammen 14 Töchter. Davon sind 5 äußerlich normale Frauen unverheiratet und deswegen nicht zu beurteilen, die 9 übriggebliebenen Töchter, haben ihre Konduktoreigen-

schaft dadurch bewiesen, daß sie entweder selber Nystagmus zeigten, oder Nystagmus-Nachkommen erhielten.

Während nun in Japan maximal pro Familie nur 2 Töchter einwandfrei als Konduktoren erkannt wurden, finde ich im von *Hemmes* bearbeiteten und im neueren Material schon einige Male 3 Konduktor-Töchter (*Burton* Abb. 8, *Auden*, *Radloff*, *Dubois* Abb. 9, *Kibort*, *Waardenburg* [Stammbaum FF Abb. 10], *Aguilar* und ebenfalls schon viermal 4 Konduktor-Töchter [*Waardenburg* Stammbaum BB], *Hemmes* [Stammbaum JJ und LL], *Mac Gregor* Abb. 11). In einem von mir hier zum erstenmal veröffentlichten Stammbaum aus längstvergangenen Jahren (Abb. 12) und im Stammbaum LL von *Hemmes* (Abb. 13) umfaßten diese vier alle Töchter, in den Stammbäumen BB und JJ je 4 von 6 Töchtern. Im Stammbaum R von *Kibort* (Abb. 14) haben einmal alle 5 Töchter eines befallenen Mannes wieder Nystagmus. Auch im Stammbaum von *Cox* (Abb. 15) erwiesen sich alle 5 Töchter aus 2 Ehen als belastet. Im Stammbaum JJ kommt sogar eine Familie vor, wo *Hemmes* 6 von 7 Töchtern selber befallen fand und wo die eine nicht-befallene Tochter unverheiratet war.

Es ist deshalb sehr wertvoll, daß es mir bei der Sippe JJ von *Hemmes* nachzuweisen gelang, daß in der Tat auch diese jüngste Tochter Konduktorin war (Abb. 16). Es ist anzunehmen, obwohl es nicht erwiesen ist, daß in den übrigen Familien phänotypisch normale Töchter, die keine Nachkommen haben, ebenfalls belastet sind.

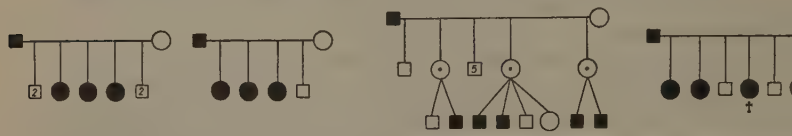


Fig. 8. Burton [1895]. *Fig. 9. Dubois* [1913]. *Fig. 10. Waardenburg, Hemmes*, *Fig. 11. Mac Gregor F. F.* [1924].

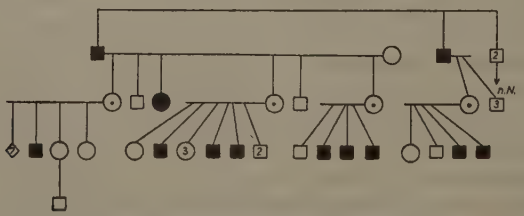


Fig. 12. Waardenburg. Fam. v. d. H.

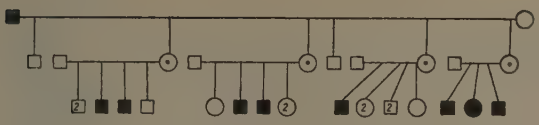


Fig. 13. Hemmes [1924]. L. L.

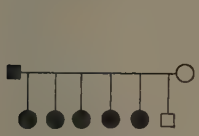


Fig. 14. Kibort [1910], *Semadeni* [1939].

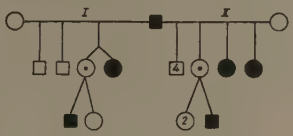


Fig. 15. R. A. Cox [1936].

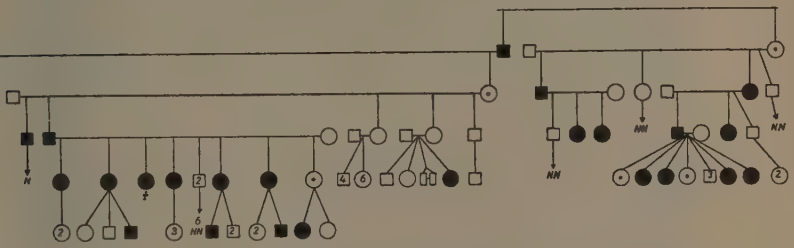


Fig. 16. Hemmes, ergänzt von *Waardenburg*.

3. Mein dritter Punkt lautete, daß Konduktorinnen ihre Nystagmusanlage auf die Hälfte der Söhne und der Töchter vererben. – Diesbezüglich ergibt sich folgendes:

Latent befallene Frauen Erdteil	Anzahl	Anzahl der normalen Söhne	befallenen Söhne	normalen Töchter	sicheren Konduktortöchter	befallenen Töchter
Japan	9	7	22	12	2	3
Europa und Amerika	92	76	131	131	29	32
Manifest befallene Frauen						
Japan	2	4	2			2
Europa und Amerika	28	22	19	31	4	15
Alles zusammen . .	131	109	176	174	35	52
		(61,75±2,88%)			87	
Anzahl der Söhne: 285			Anzahl der Töchter: 216			

In Japan sind erstens zuviel Söhne geboren (35 gegen 19 Töchter) und von diesen sind die befallenen Söhne wieder signifikant in der Mehrzahl (24 gegen 11 normale Söhne). Das bleibt so, auch wenn man die 6 normalen Söhne und 4 normalen Töchter der drei nicht-befallenen Schwestern in dieser Familie hinzuzählt. Man erhält dann ein Verhältnis von 17 normalen zu 24 befallenen Söhnen. Für die Töchter gilt, daß 12 normale (teils Konduktor) 7 sicher belasteten gegenüberstehen. Die Zahlen sind jedoch für weitgehende Schlüsse zu gering.

In Europa und Amerika kommt in der allgemeinen Sexualproportion keine Abweichung von der Regel vor (250 Männer : 242 Frauen). Die Anzahl der befallenen Söhne (152) jedoch übersteigt diejenige der normalen Söhne (98) signifikant.

Das Verhalten bei den Konduktorinnen ist nicht zu beurteilen, da eine unwillkürliche Auslesewirkung entsteht, solange wir nur von Fällen mit befallenen Söhnen ausgehen. In diesem Material fehlen dann diejenigen Konduktorinnen, die zufälligerweise keine befallenen Söhne haben und die, wenn sie berücksichtigt waren, dadurch zur Kompensation des Übermaßes von befallenen Söhnen beigetragen hätten.

Um den Fehler zu vermeiden, die richtige Zahl der Konduktorinnen, die nicht bekannt ist, abzuschätzen, habe ich zum Studium der Kinder von Konduktorinnen nur alle Töchter befallener Väter genommen, da diese bei X-chromosomalem Erbgang sicher Überträgerinnen sind. Dann stellt sich folgendes heraus:

Autoren	Normale Söhne	Befallene Söhne	Normale Töchter	Latente Konduktorinnen Töchter	Manifeste Konduktorinnen Töchter
Mikamo	1	1			
Toyada	1	7	4		
Kitahara (Fam. 2) .	4				1
Kitahara (Fam. 1) .	1	2	6		1
Waardenburg (DD) .	3	2			
Waardenburg (FF) .	3	7	3		
Hemmes (HH) . . .	2	2	3		
Hemmes (JJ) . . .	4	5	9		1
Hemmes (LL) . . .	5	9	6		
Audeaud	2	2			1
Mac Gillivray . . .	5	2	6	2	1
Gunn		4	3		
Auden	4	5	8		
Nodop	2	4	8		
Radloff		1			
Dubois	2	4		2	
Zusammen	39	57	57	7	5

Die Ziffern für die Töchter sind unbrauchbar, da wir wieder nicht wissen, wieviele Überträgerinnen sich unter den sogenannt normalen Töchtern befinden. Von den Söhnen sind 59,37 $\pm$ 5,012 % befallen und 40,63 $\pm$ 5,012 % normal, was innerhalb des Zufalls liegt. Deshalb ist die Schlußfolgerung, die ich im Jahre 1935 gezogen habe, daß beim extraokularen Nystagmus das belastete Chromosom eine gewisse Prävalenz besitzt, mathematisch nicht mehr haltbar. Ich fand nämlich für die damals zur Verfügung stehenden Sippen des Schrifttums bei Söhnen von Konduktorinnen 66,2 $\pm$ 2,93 Prozent befallen, was sich nicht sehr von den jetzt aus den seitdem erschienenen Daten mitberechneten 61,75 $\pm$ 2,88 Prozent unterscheidet. Dieses Material umfaßte jedoch nicht alle sicheren Konduktorinnen in den Geschwisterreihen samt ihren Söhnen. Da die Zahlen klein sind, bleibt es wünschenswert, diese Frage in Zukunft an größerem Material zu lösen.

Summary.

A critical review is given of the whole literature on hereditary extraocular nystagmus. The majority of cases shows an irregularly dominant or a recessive X-chromosomal inheritance. The author however assumed, also from more recent cases in the literature, the existence of dominant and recessive autosomally inherited cases. The question whether there are significantly more affected than non-affected brothers can only be solved by enlarged material. The author presents some new experiences of his in this field.

Résumé.

Une étude critique de toute la littérature sur le nystagmus extraoculaire héréditaire rend probable que la plupart des cas se transmettent d'une façon irrégulière dominante ou récessive X-chromosomale. Mais dans la littérature contemporaine également des cas ressortent qu'on ne peut expliquer que par la transmission de tares autosomales dominantes ou récessives. Une augmentation des cas X-chromosomals sera nécessaire pour décider la question s'il y a un surplus significatif de frères atteints sur les frères non-atteints. Présentation d'expériences nouvelles sur ce sujet.

Zusammenfassung.

Nach kritischer Prüfung des Gesamtschrifttums über erblichen extraokularen Nystagmus scheinen die meisten Fälle unregelmäßig-dominant bzw. rezessiv X-chromosomal vererbt zu werden. Es bleiben jedoch Fälle übrig, auch in der neueren Literatur, die vorläufig nur dominant oder rezessiv autosomal zu erklären sind. Bei den X-chromosomalen Fällen ist zur Entscheidung der Frage, ob signifikant mehr befallene Brüder als normale anwesend sind, mehr Material notwendig. Vorführung neuer, eigener Befunde.

LITERATUR

Für Autoren vor 1932 sei verwiesen auf

Waardenburg, P. J.: Das menschliche Auge und seine Erbanlagen, Nijhoff, Den Haag 1932. – *Aguilar, A.:* Sem. med., B. Aires *1*, 72, 1937. – *Allen, M.:* J. Hered. *33*, 454, 1942. – *Barrigozzi, C.:* Rass. ital. Ottal. *3*, 733, 1934. – *Bhadurio:* Calcutta med. J. *26*, 311, 1932; Ref. Amer. J. Ophthal. *15*, 984, 1932. – *Cox, R. A.:* Arch. Ophthal., Chicago *15*, 1032, 1936. – *Cuendet, J. F.* and *V. Della Porta:* Ophthalmologica *117*, 199, 1949. – Arch. Ophthal. *10*, 187, 1950; *Fattovich, G.:* Giorn. Psichiat. *64*, 393, 1936. – *Glover, L. P.:* Arch. Ophthal., Chicago *17*, 705, 1937. – *Käser, O.:* Inaug. Diss. Zürich 1941; Schweiz. med. Wschr. *31*, 846, 1942. – *Knighton, W. S.:* Arch. Ophthal, Chicago *2*, 437, 1929. – *Komai, T.:* Pedigrees of Hered. Diseases and Abnormalities Found in the Japanese Race, Kyoto 1934. – *Leroy Billings, M.:* J. Hered. *33*, 457, 1942. – *Mac Gregor, I. S.:* Ann. Eug. *13*, 135, 1946. – *Rucker, C. Wilbur:* Amer. Ophthal. *29*, 1534, 1946; Amer J. hum. Genet. *1*, 52, 1949. – *Semadeni, B.:* Schweiz. med. Wschr. *44*, 1077, 1939. – *Waardenburg, P. J.:* Z. indukt. Abstamm.-Vererblehre *70*, 358, 1935; Ned. Tijdschr. Geneesk. *80*, 49, 5523, 1936; Docum. ophthal. *4*, 166, 1950.

Wrighton, R. F.: Acta genet. *4*, 312–343, 1953

THE THEORETICAL BASIS OF THE THERAPEUTIC TRIAL

By R. F. WRIGHTON

Department of Medical Statistics, University of Birmingham
and Department of Social and Industrial Medicine, University of Sheffield

1. Introduction.

There has been little fundamental change in the concepts underlying the therapeutic trial since public controversy gave rise to a detailed discussion by *Greenwood* and *Yule* [1915] of the statistics of anti-typhoid and anti-cholera inoculations. The canons they established comprised the commonsense notion that there must exist treated and untreated groups which are comparable in all

relevant respects, together with the more sophisticated requirement that the acceptance of an observed difference in the cure- or attack-rates must be qualified by a consideration of the magnitude of difference which might be expected to arise on the assumption of equal rates and a purely chance fluctuation[1]. The χ^2-test elaborated by *Pearson* at the beginning of the century was invoked with the intention of providing a suitable disciplinary procedure and its use in a modified form has continued to the present day as the most common statistical technique employed in the therapeutic trial.

It would be difficult to disentangle the conceptions and misconceptions which ultimately gave rise either to this particular application of *Pearson*'s test (the first in the field under consideration), or to *Pearson*'s own investigation, or to the widespread recognition of the test of significance as a weapon of inference. We may note, however, that the basic idea has a long history. Before 1900 several mathematicians had used the concept of randomness to provide an empty hypothesis against which to carry on heuristic discussion of causes. We may cite *Daniel Bernoulli's* attempt to demonstrate the causal origin of the clustering of the planes of the planetary orbits around the ecliptic, *Kirchhoff's* argument for identifying the dark lines in the solar spectrum with the absorption bands of terrestrial substances and *Michell's* discussion of the possibility of the purely optical origin of the phenomenon of binary stars[2]. Later, in a field more close to our own, Lexis used a fundamentally similar notion to detect the existence of phenomena underlying the irregularities of social statistics (see *Keynes* [1921]). Finally, *Pearson* himself appears to have developed the χ^2-test primarily for the purpose of testing the goodness of fit of hypothetical curves to frequency distributions of assumed stochastic origin.

Whatever the credentials of the method of the significance test in these situations, the reasoning by which the concepts involved are transferred to the domain of the therapeutic trial does not become more clear on reflection. Two circumstances may perhaps have some bearing on the popularity of the cult. In the first place, an undercurrent of academic scepticism appears to have led many statisticians seriously to believe that even in the conduct of practical affairs a hypothesis must be supposed to be susceptible only of disproof if

[1] For an early discussion see chapter 3 of *Gavarret*'s Statistique Médicale (1840).

[2] See *Jevons*, Principles of Science (1874).

logical rectitude is to preserved[1]. Secondly, the statistical method has been generally regarded as fundamentally different from the established method of the experimental sciences; and historically this is seen to be true if we recognise that its advanced techniques were primarily developed over a long period with an eye to the analysis of bodies of data with extent and nature determined, either wholly or in part, independently of the eventual manipulator. This is particularly noticeable in the attitude of *Pearson*, in his time the principal exponent of the analytical approach to statistical inference. His applications of the goodness of fit test strikingly reflect this outlook[2].

With these doubts in mind, *Hogben* and *Wrighton* [1952] have provisionally attempted to examine the fallacies and weaknesses inherent in the customary procedures of statistical analysis in the therapeutical trial and in the alternatives that are available in statistical literature. It is the purpose of the present paper to re-define and to re-examine the problem. Since it will be convenient to introduce new nomenclature for some concepts which are already current it will not be out of place to give a brief review of the revaluation of the credentials of statistical procedures which has been in progress during the last twenty years.

2. *The concept of probability.*

Since the publication of *Bayes*' Essay and more particularly as a result of the appearance of *Laplace's Mémoire* in 1774, there has existed a confusion between two entirely different meanings of the

[1] Cf. *Fisher* [1951] p. 16: "Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis."

[2] See, for example, *Pearson* [1900, 1911], especially the latter where he suggests an application relevant in the present context. We cite his actual words: "We have two records of the number of rooms in houses where (i) a case of cancer has occurred, (ii) a case of tuberculosis has occurred; the number of cases of each disease may be quite different, and we may not be acquainted with the frequency distribution of the number of rooms in the given district. What is the chance that there is a significant difference in the tuberculosis and the cancer houses? Or again, we have a frequency distribution of the interval in days between bite and onset of rabies in two populations of bitten persons (i) who have been and (ii) who have not been inoculated in the interval. What is the probability that the inoculation has modified the interval?" Leaving aside the question of what precisely is meant by *the chance that there is a significant difference*, we may note that in each case which *Pearson* gives it is proposed to invoke a statistical device in order to argue from an established body of data whose extent is presumably arbitrary relative to the needs of the investigator.

word *probability*. *Carnap* [1950 and elsewhere] conveniently refers to these as *probability*$_1$ and *probability*$_2$. Briefly, and somewhat elliptically, probability$_2$ describes physical phenomena which involve randomly-occurring events; probability$_1$ is relevant to any attempt to formulate a logical calculus concerned with degrees of credibility, degrees of confirmation or other entities which the exponents of such theories might regard as having their expression in the *mind*.

We shall avoid confusion if we make clear at the outset that we shall be concerned exclusively with probability$_2$. In particular, phrases such as *the probability that treatment A is better than treatment B* or *the probability that the cure rate of treatment A lies between* p_1 *and* p_2 will be regarded as strictly meaningless within the vocabulary we employ. We are however, by no means on this account released from the embarrassment of treading on ground which remains highly controversial. There remain in fact two broad questions for fundamental discussion:

(i) what criteria must we adopt in adjudging a phenomenon as random?
(ii) if we grant the assumption of randomness, what elaboration of a formal theory is relevant to practical affairs?

With regard to the second of these we shall maintain that we do not absolve ourselves from discussion of a class of situations, treated by the classical writers on *inverse probability*, situations which some contemporary writers would imply have been conveniently sidestepped by the modern development of the theory of continuous probability functions, when we abandon the notion of *probability in the mind*. The most elementary of these problems concerns assessment on the basis of repeated trials of whether a coin has been falsely minted with two heads. We shall here regard this as typical of the class of situations with which a theory of statistical inference should properly deal. The Laplacian method for dealing with such problems is widely recognised as having been founded on error. We here maintain, again following *Carnap*'s dichotomy, that its basic fallacy, in its more plausible applications, lay in asking questions of situations defined within the physical framework of probability$_2$ in such a way that answers had to be given in terms of probability$_1$. The type of question which must be asked of a theory, if a consistent vocabulary is to be employed, has become increasingly manifest in the last decade.

3. The current revaluation of the credentials of statistical procedures.

Discussion of the conceptual basis of statistical theory since 1930 has focussed on the second of the questions propounded above; that is to say, its main concern has been with what particular elaboration of a formal theory is of practical utility. It is significant that the most important clarifications have taken place in the applied domains of agricultural economics and industrial inspection. The basic change of emphasis, as suggested above, has consisted in discarding the conception of mathematical statistics as a tool for the *analysis of data*. Instead, it is required that the theory should provide *a priori* rules for the conduct of statistical experimentation.

The first fully explicit recognition of this principle seems to have been made by *Neyman* [1938] in an exposition of his theory of *confidence intervals*. In the same paper however, as in his earlier contributions, *Neyman* in discussing *best systems* of confidence intervals develops his theory in a way which, to the writer, seems inconsistent with a full acceptance of the implication of the phrase *inductive behaviour* there introduced. In order to avoid the appearance of an implicit acceptance of the whole *Neyman*'s theory, no use has been made in what follows of the term *confidence intervals*. In an attempt to reconcile the theories of different statistical schools, *Kendall* [1949] has raised objections to *Neyman's* theory, objections which seem to the writer to be valid and to spring from a similar appreciation of the inconsistency of *Neyman*'s incomplete acceptance of the necessity for a reorientation.

Neyman's ideas have received a more elaborate development in the work of the late *Abraham Wald*, especially in his theory of statistical decision functions (*Wald* [1950]) where there is to be seen a complete break with the older analytical outlook. *Wald's* theory, however, has been largely developed against the background of risks definable in monetary terms. Indeed, it has up to now claimed its principal application in the field of industrial inspection and experimentation. There, in addition, the earlier work of *Shewhart* has led to a peculiar interpretation attaching to the empirical basis of the requisite postulates of underlying randomness. It is the writer's view that our problem in this context is different. In the therapeutic trial our primary concern is to make statements and not, as in private commercial applications, to arrive at decisions; the *risk* associated with making false statements will not be commensurable

with the *loss* incurred by a profligate use of experimental material. Moreover, the foundation of a properly applied stochastic theory in a process which may justifiably be regarded as random requires examination in a fashion differing greatly from what is necessary in a quality control situation. Accordingly, while taking over *Wald's* concept of *terminal decisions* we shall refer to *terminal statements* and in other respects develop our approach independently of the theory of statistical decision functions.

4. The introduction of probability theory into the therapeutic trial

A statistical methodology does not underly all biological experimentation; and it would be ludicrous to assert this in the absence of extreme statements to the opposite effect which may be met in the literature. We shall accordingly define at the outset what logical dilemma it is that demands in the particular circumstances of a therapeutic trial an approach which differs from the commonsense attack of the universally-accepted experimental procedure.

We suppose that the value of a new treatment (treatment B) is to be measured against that of the customary treatment (treatment A) and we make the simplifying assumption that the outcome of treatment of the individual is measured by whether or not he is, by some recognised standard, cured or not cured after a fixed period of time. The basic requirement that we shall postulate in all that follows is that *the outcome of the disease in the individual under treatment A shall not be predictable with assurance.* This principle will be regarded as necessary in justifying the conduct of a trial at a statistical level.

In these circumstances we should like to have grounds for saying of a subject whether or not he possesses the compound property which implies that *he will not be cured if given treatment A but will be cured if given treatment B.* That this is impossible in the type of situation we have postulated (a situation which will be more generally relevant in the domain of prophylaxis than in therapeutics) is the fundamental dilemma which makes a statistical approach necessary.

It may happen that the estimate of efficacy must be made on the basis of a comparison of records made before and after the introduction of a new treatment. From the standpoint adopted in the present paper this comparison will be regarded as properly made if based on an empirically acquired knowledge of the fluctuation of analogous cure rates rather than by means of an appeal to any formal

statistical calculus. Probabilistic considerations will then enter, if at all, at a level inconsistent with a high degree of formalisation; and doubt will legitimately attach to any rigid interpretation of the results.

The controlled trial will be more satisfactory if only because cure-rates referable to a given treatment vary as the result of changes in ancillary methods. It is for this procedure that we seek an adequate rationale. The simplest form of trial will be that in which the N trial subjects are divided into two groups, m being subjected to treatment A and n to treatment B. Although the classification has no meaning referable to direct inspection, we may suppose each individual to fall into one of four classes according to his potential reaction to each of the treatments. We suppose the numbers in these classes to be α, β, γ, δ, as follows ($\alpha+\beta+\gamma+\delta = m+n = N$).

		Treatment B	
		Curable	Not Curable
Treatment A	Curable	α	δ
	Not Curable	β	γ

As a result of the trial it is hoped to make some form of statement or statements about the proportions $\alpha/(m+n)$, $\beta/(m+n)$, $\gamma/(m+n)$, $\delta/(m+n)$ or about their relative values. Intuitively we see that, m and n being sufficiently large, we may hope to estimate $\frac{\alpha+\beta}{N}$ and $\frac{\alpha+\delta}{N}$ (or what is equivalent, $\frac{\gamma+\delta}{N}$ and $\frac{\beta+\gamma}{N}$) with a high degree of accuracy. From this approach, at no stage do we see ourselves in a position to make statements purporting to have relevance outside the group of subjects actually concerned in the trial.

We provide a sufficient basis for the legitimate application of probability theory by dividing the subjects into two groups by a randomising process. Perfect randomisation is not attainable. It is sufficient that the process should be adequately random for the purpose in hand; and on the basis of accumulated evidence it seems reasonable to grant that such a degree of randomisation is achievable by mechanical means[1]. It is important, however, to stress that with

[1] In this respect we may regard as a *recipe Poincaré*'s *definition* of probability as the consequence of small causes giving rise to disproportionately large effects.

the above formulation we may carry out dummy trials, in which α, β, γ, δ are known, alongside of the real trial and equivalent to it in every relevant respect. By doing this as many times as we like with arbitrary combinations of α, β, γ, δ we may verify with a certainty analogous to that obtainable in non-statistical contexts whether or not the device we are using tends to lead to results which accord with its pretentions.

We may summarise the essence of the last few paragraphs by saying that we need not introduce the theory of probability into our analysis of the controlled therapeutic trial by reliance on an assumption, almost certainly gratuitous and never verifiable, that patients who differ with respects to potential response to treatment will submit themselves for treatment in random order. Some analogous but loosely formulated assumption may well be required if the results of the trial are to have practical value; but as an exercise in a rigorous method, we may with profit restrict our inferences to the group of subjects directly enumerable by the investigator. The strictly classical calculus of probability, in particular the part concerned with urn models, will now be adequate for a formal theory. If we concern ourselves with a continuum of probabilities it will be purely as a matter of computational convenience.

The apparatus of the relevant urn-model is then as follows: we have an urn containing balls, complementary hemispheres of which we label explicitly as A and B.

α are white on both hemispheres.
β are white on the hemisphere labelled B and black on that labelled A.
γ are white on the hemisphere labelled A and black on that labelled B.
δ are black on both hemispheres.

All sampling is done without replacement and the colour of only one hemisphere may be observed in an individual drawing.

It will help us to formulate the proper place of probability as an instrument of inductive reasoning in the type of situation we have indicated if we first examine some situations in which no uncertainty attributable to sampling enters into our final verdict.

5. *Non-statistical inference.*

In the simplest case we suppose that we have a number of hypotheses $H_1 \ldots H_n$, one of which we may suppose *a priori* to be true. These we shall refer to as the *prior hypotheses*. An experiment is planned which may have as its outcome one of the observational

results $O_1 \ldots O_k$. We may then set out a table in which there is a row of cells corresponding to each hypothesis and a column to each possible observation. We require that it should be possible to mark in each row by deductive means the observations which may follow from the corresponding hypothesis. We shall refer to an experiment which admits of such a specification as *holonomic*. The experiment is performed and the observation O_i is recorded. We may then assert that one of those hypotheses is true whose place in the i^{th} column is marked.

As an example of a situation in which such a procedure may be illuminating we may borrow a classical problem from the epistemologist. A paradox is said to have been propounded by *Epimenides* the Cretan who asserted that all Cretans were liars[1]. In conformity with the custom of text-books of logic we interpret the word *liar* to signify someone who never tells the truth and a person who is not a liar as one who is consistently truthful. We may accordingly regard the statement of *Epimenides* as the observation in a holonomic experiment of which the prior hypotheses are:

A. All Cretans are liars.
B. No Cretans are liars.
C. Some are and some are not; *Epimenides* is.
D. Some are and some are not; *Epimenides* is not.

To conceive the problem as an experiment we postulate that we give *Epimenides* the opportunity to make a single affirmation corresponding to any one of the above. In the text-book example the outcome is the observation A. We may construct the observation-hypothesis table for such an experiment as follows:

Hypothesis	Epimenides Asserts			
	A	B	C	D
A		+	+	+
B		+		
C	+	+		+
D				+

Within the framework of our four admissible hypotheses we then deduce from the first column that, if the outcome of the experiment is that the recorded observation is *A*, hypothesis *C* is true. We thus see that we could not have felt sure of obtaining an unambiguous

[1] See Epistle of Paul to Titus I. xii.

decision before the performance of the experiment. With this example in mind we may distinguish two particular types of experiment:

(i) If, in the event, the observation O_i is made, it may occur that in the i^{th} column there lies only one marked cell. We refer to this experiment as *adequate a posteriori*.

(ii) If every column contains only one marked cell we refer to the experiment as *adequate a priori* relative to the prior hypotheses.

A useful experiment need not necessarily fall in the second class.

We consider now a more complicated case. If, for example, hypotheses and observations are numerical measures, we may neither wish nor be able to assert the truth of one particular hypothesis. If we confine ourselves to discrete variables, using $x = 1, 2, 3 \ldots$ for the observations and $X = 1, 2, 3 \ldots$ for the hypotheses we may characterise the experiment by a table which in a particular case might be as follows:

X \ x		1	2	3	4	5	6	7	8	9	10	11
		O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}	O_{11}
1	H_1							+	+	+	+	+
2	H_2							+	+	+	+	
3	H_3						+	+	+	+		
4	H_4					+	+	+	+			
5	H_5				+	+	+	+				
6	H_6			+	+	+						
7	H_7			+	+	+						
8	H_8		+	+	+							
9	H_9	+	+	+								
10	H_{10}	+	+									

Here, given the observation O_4, we must accept one of the hypotheses $H_5 \ldots . H_8$, that is to say we may assert $5 \leqslant X \leqslant 8$. In a case like this we need to modify the concept of adequacy.

We may regard the experiment as having been justified if one of a certain set of sets of hypotheses can be asserted to be true as a result of it. This will comprise what we shall refer to as the *set of acceptable terminal statements* and, *relative to this set*, we may then refer to an experiment, as before, as being *a priori* or *a posteriori* adequate. For example, relative to the set of acceptable terminal statements $n \leqslant X \leqslant n+4$, $n = 1 \ldots 6$, the experiment schematised

above is *a priori* adequate; relative to the terminal statements $n \leqslant X \leqslant n+3$ and the observation O_6 it is *a posteriori* but not *a priori* adequate.

If O_1 is observed we may assert that $9 \leqslant X \leqslant 10$. The result is unnecessarily selective relative to both of the sets of acceptable terminal statements we have put forward. In general we shall obtain a greater compactness in an experiment-design if we confine ourselves to sets of acceptable terminal statements which are *minimal*, that is to say sets in which:

(a) all members are different.
(b) no one implies any other statement which is also acceptable.
(c) no statement includes an element outside the range of prior hypotheses.

We have introduced six concepts, viz. those of:

(i) the prior hypothesis
(ii) the holonomic experiment
(iii) the acceptable terminal statement
(iv) *a posteriori* adequacy
(v) *a priori* adequacy
(vi) the minimal set of acceptable terminal statements

Only one further definition is required in the statistical theory of the therapeutic trial to which we now turn. An element of uncertainty there arises and our purpose in making this digression has been to stress the independence the *majority* of concepts required in the theory have to stochastic considerations.

6. Single stage statistical inference.

With the introduction of sampling procedures we encounter a factor which calls for an extension of our definition of a holonomic experiment. We shall refer to an experiment as *holonomic in the statistical domain* if, as before, it is possible to cite prior hypotheses and observations possible on these hypotheses; but instead of assigning equal weights to each observation consistent with the hypothesis H_i, we shall require to specify each by a calculable probability-value, the sum of all such being unity. Within this framework we now seek to extend the logical procedure which has been followed in non-stochastic experiments.

In general, no application of the theory of probability is able to do more than assign a probability differing from one or zero to the occurrence of an event; and no event of ultimate practical importance will be usefully predictable within the theory unless this probability is acceptably high. Implicitly, as in some physical applications, or

explicitly, as in the application of *Bernoulli*'s theorem to games of chance[1], we must accordingly and always work within the limitations of a *theoretical uncertainty level* defined as ε, the theoretical probability of the occurrence of an event or group of events deemed to be useful being $\geqslant 1-\varepsilon$. In the context of statistical experimentation the aim of a theory will be to make statements that are *almost certainly true*, in the sense that the theoretical probability of their falsehood is less than or equal to the preassigned value of ε. In the case of a holonomic experiment this will mean that the probability of making a true statement is $\geqslant 1-\varepsilon$, relative to the assumed truth of *each* prior hypothesis taken individually.

If we fix in advance our uncertainty level we may in each row mark off observations the sum of whose probabilities on the relevant hypothesis is $\geqslant 1-\varepsilon$. If we treat the table in the manner described for non-statistical experiments, it follows directly that we shall have ensured the validity of the assertion that one or other of the marked hypotheses in the relevant observation-column is true with probability $\geqslant 1-\varepsilon$.[2] This is the basic result needed. Its ready derivation may not indicate its true position as the corner-stone of an inverse statistical calculus.

A fundamentally new element has been introduced into experiment-theory by this extension in that it allows us a choice with respect to the marking of the rows, a choice which has no analogue in the non-statistical case. We shall need to examine the extent of this apparent arbitrariness and its implications; and we shall maintain as a result of this examination that two conditions must be satisfied by a holonomic experiment in the statistical domain.

(i) The experiment should be *a priori* adequate relative to the set of acceptable terminal statements.

(ii) The set of acceptable terminal statements should be *minimal*.

By definition, a set of acceptable terminal statements will comprise statements any one of which we may regard as having

[1] Cf. *Uspensky* [1937] pp. 103–109. *Bernoulli*'s theorem preassigns a theoretical uncertainty level and a desired minimum gain and evaluates the number of games that must be played in order that this minimum gain shall *almost certainly* accrue to that player whom the rules favour. There are obvious analogies with the inverse type of problem discussed later.

[2] It will save excessive circumlocution if we use this phrase, which taken literally is meaningless within the vocabulary we have postulated, to connote the idea that the statement is made after an experiment designed in such a way that true statement must result from it with probability $\geqslant 1-\varepsilon$.

justified the performance of the experiment from which it results. We may say alternatively that they will all conform to standards of precision which would merit the incorporation of any one of them, if true, into an accepted body of knowledge. We sharpen the outlines of the situation if we demand in what follows that the standards pertaining to this body of knowledge should be fixed and unequivocal. That is to say, all statements resulting from a statistical experiment will be strictly regarded as worthless if they are not acceptable.

Consider the following example. We are given two coins either or both of which may have been minted with two heads or two tails. An experiment is set up in which each coin is tossed four times. The *total* number of heads obtained as a result is recorded. The schema of the experiment is then as in the table.

Hypotheses		Observations Expressed as No. of Heads Recorded								
		0	1	2	3	4	5	6	7	8
A	TT, TT	**1**	—	—	—	—	—	—	—	—
B	TT, HT	**$\frac{1}{16}$**	**$\frac{4}{16}$**	**$\frac{6}{16}$**	**$\frac{4}{16}$**	**$\frac{1}{16}$**	—	—	—	—
C	TT, HH	—	—	—	—	**1**	—	—	—	—
D	HT, HT	**$\frac{1}{256}$**	**$\frac{8}{256}$**	**$\frac{28}{256}$**	**$\frac{56}{256}$**	**$\frac{70}{256}$**	**$\frac{56}{256}$**	**$\frac{28}{256}$**	**$\frac{8}{256}$**	**$\frac{1}{256}$**
E	HH, HT	—	—	—	—	**$\frac{1}{16}$**	**$\frac{4}{16}$**	**$\frac{6}{16}$**	**$\frac{4}{16}$**	**$\frac{1}{16}$**
F	HH, HH	—	—	—	—	—	—	—	—	**1**

Let us suppose that we are allowed to work within the bounds of an uncertainty level of $\varepsilon = {}^{1}/_{12}$; then the marking indicated in the table by heavy type represents *one* way in which an experimental rule may be set up. If the object of the trial is to assess the number of heads distributed between the two coins we may summarise the operation of the rule as follows:

Heads Observed	0	1	2	3	4	5	6	7	8
Assertion	0	1	1 or 2	1 or 2	1, 2, 3	2 or 3	2 or 3	2 or 3	4

Let us now suppose that of the statement possible on the basis of this, the one corresponding to the observation of four heads would not be admissible into the hypothetical body of knowledge. We examine the consequences of modifying the rule so that such observations are ignored. The schema becomes:

Hypotheses		Observations, Expressed as No. of Heads Recorded							
		0	1	2	3	5	6	7	8
A	TT, TT	1	—	—	—	—	—	—	—
B	TT, HT	1/15	4/15	6/15	4/15	—	—	—	—
C	TT, HH	—	—	—	—	—	—	—	—
D	HT, HT	1/186	8/186	28/186	56/186	56/186	28/186	8/186	1/186
E	HH, HT	—	—	—	—	4/15	6/15	4/15	1/15
F	HH, HH	—	—	—	—	—	—	—	1

We may see more clearly the possible effect on the uncertainty level and on the general set-up if we examine separately the outcome if successive hypotheses exclusively hold.

If A or F holds we shall, as before, be consistently correct in our assertions. If B or E hold we shall make false statements one in fifteen times; and this is consistent with $\varepsilon = {}^{1}/_{12}$. If, however, D holds we shall be wrong ${}^{18}/_{186}$ times in the long run and the uncertainty level is vitiated. Worse still, if C holds we shall invariably be wrong in our assertions.

The reason for the first requirement above is now obvious. In general, and provided we take seriously the concept of a *corpus of knowledge*, the uncertainty level is likely to be vitiated by the exclusion of any class of experimental result. In other words, an experiment must be *a priori adequate* if the validity of the uncertainty level is to be retainable[1].

Consider now the same model situation but with uncertainty level ${}^{3}/_{8}$. The following abbreviated schema will be seen to provide a rule which is formally correct.

	0	1	2	3	4	5	6	7	8
A	+	—	—	—	—	—	—	—	—
B	—	—	+	+	—	—	—	—	—
C	—	—	—	—	+	—	—	—	—
D	—	—	—	+	+	+	—	—	—
E	—	—	—	—	—	+	+	—	—
F	—	—	—	—	—	—	—	—	+

[1] On going to press the writer notes that this point is made in a special context by *Neyman* (Lectures and Conferences on Mathematical Statistics and Probability, 1952, p. 257).

The corresponding summary of the rule will follow as:

Heads Observed	0	1	2	3	4	5	6	7	8
Assertion	0	?	1	1 or 2	2	2 or 3	3	?	4

The query under the observation of one or seven heads may be interpreted as implying that in these circumstances a statement must be made which is known to be false; that is to say, it must be asserted that none of the prior hypotheses is true. The chance of this occurring in similar experiments in which the uncertainty level is given a more realistic value is remote; its probability will certainly be less than the uncertainty level. This ensures that in *ignoring* such results the uncertainty level will not be vitiated. The difficulty is therefore of a different order of importance from the first. But there remains a clumsiness in the theory which will become more apparent later when we examine the therapeutic trial model in a similar way. We note in addition that, laying aside the inconveniences of the queried entries in the last table, those entries remaining are of unequal precision.

Both of the difficulties mentioned above are avoidable if we demand initially that the list of acceptable terminal statements should be minimal according to the definition of p. 322. This may appear as an unjustifiable limitation of the scope of the theory in that the experimenter will thereby be prevented from gaining any advantage from a fortuitous increase in accuracy. If however we formalise the situation in terms of the vitally important concept of a hypothetical body of knowledge we shall see that any such advantage is illusory. It presupposes that some statements which may result from the experiment may be incorporated into a body of knowledge possessing higher standards of precision than those originally postulated. If this is done then *relative to this second body of knowledge*, the experiment is *a posteriori* but not *a priori* adequate. The unfortunate consequences of such a design have been sufficiently emphasised. It follows that no advantage is to be gained by using a set of terminal statements which is not minimal; moreover such a custom will facilitate a relapse into the acceptance of a highly fallacious instrument of experimentation.

We are now able to state and solve an adequately defined problem in statistical inference. For the above model let us suppose

that the minimal set of terminal statements comprises the assertions that the number of heads is:

0 or 1; 1 or 2; 2 or 3; 3 or 4.

With an uncertainty level of $^1/_{16}$, four tossings of each coin are then sufficient for a satisfactory experiment and a suitable abbreviated schema will be as follows. It is by no means the only one that adequately fulfills the intended purpose.

	0	1	2	3	4	5	6	7	8
A	+	—	—	—	—	—	—	—	—
B	+	+	+	+		—	—	—	—
C	—	—	—	—	+	—	—	—	—
D		+	+	+	+	+	+	+	
E	—	—	—	—	+	+	+	+	+
F	—	—	—	—	—	—	—	—	+

The corresponding summary of the rule is then

Observation	0	1	2	3	4	5	6	7	8
Assertion	0 or 1	1 or 2	1 or 2	1 or 2	2 or 3	2 or 3	2 or 3	2 or 3	3 or 4

We shall regard this as typical of a well-designed statistical experiment and we may now state formally the problem of single-stage inference for holonomic experiments in the statistical domain. We postulate initially:

(a) a model set-up comprising prior hypotheses and an allowable regimen of experimentation;

(b) an uncertainty level;

(c) a set of acceptable terminal statements which is minimal.

On this understanding we aim to design an experiment as a result of which we assert one of the acceptable terminal statements to be true within the limits imposed by the uncertainty level. In addition, we may require that the experiment designed is the most economical possible in conformity with preassigned standards. In the model situation we have used as an illustration we might thus require that the number of times each coin is tossed shall be as small as possible. Even with this stipulation, there is no reason to suppose that there is not devisable a multiplicity of suitable rules in any situation. We may remark that any choice that then exists will be completely arbitrary.

7. The problem of the single-stage therapeutic trial.

We may now precisely define the analytical problem presented by the therapeutic trial. The model situation and permitted sampling procedure have been stated; and it may be supposed that we can agree upon a value for the uncertainty level. There remain to be defined the range of prior hypotheses and the minimal set of acceptable terminal statements. These given, it will then be necessary to evaluate the number of subjects needed in the two groups in order that one of these terminal statements may be made and to define concurrently a rule of exclusion adequate to the achievement of this end. In this section, we discuss the specification of the prior hypotheses and the acceptable terminal statements and indicate a formal method of obtaining an adequate rule which is most economical, *providing an adequate rule can be supposed to exist*[1]. The method will be of little practical value. The primary object of this communication, however, is to assess at a conceptual level what part the statistical method has to play in the therapeutic trial, rather than to provide rules for its use in practice.

The choice of the prior hypotheses and acceptable terminal statements will depend entirely on the nature of the application. We may consider, for illustrative purposes, two particular classes:

(a) *Prior hypotheses:* all combinations of non-negative values of α, β, γ, δ consistent with $\alpha + \beta + \gamma + \delta = m+n$.

Minimal set of acceptable terminal statements: all interval estimates of the form $\vartheta \leqslant \frac{\beta-\delta}{m+n} \leqslant \vartheta + l$, where $-1 \leqslant \vartheta \leqslant 1-l$, and l is a constant.

(b) *Prior hypotheses:* all combinations of non-negative values of α, β, γ consistent with $\alpha + \beta + \gamma = m + n$.

Minimal set of acceptable terminal statements: all interval-estimates of the form $\vartheta \leqslant \frac{\beta}{m+n} \leqslant \vartheta + l$ where $0 \leqslant \vartheta \leqslant 1-l$ and l is a constant.

Thus, with respect to prior hypotheses, (a) is the most general case possible; (b) is the most general possible if it may be assumed that it is not conceivable that any subject who is curable by the standard treatment is not also curable by the new treatment.

Suppose we are given values for m and n. The number of ways groups of sets of excluded hypotheses consistent with the uncertainty level may be selected in the rows of the observation-hypothesis table is enumerable, though very large even for small values of m and n.

[1] A proof that such a rule does exist is given in Section 8.

Each way will yield a set of potential terminal statements. Some of these will take the form of interval-estimates of $\frac{\beta-\delta}{m+n}$. The majority will not. The uncertainty level is not vitiated if we extend such statements by adding to them those prior hypotheses which will convert them into intervals of the minimum possible length. Corresponding to each of the resulting ways of selecting a rule we shall then have a set of interval-estimates of $\frac{\beta-\delta}{m+n}$. Denote these by $S_1, S_2 \ldots S_r$ and let the length of the largest interval in the set S_1 be l_1. Now let $l_{m,n}$ be the minimum value of the l_i among all the S_i and denote by $\Sigma_{m,n}$ the group of rules which have this value for l_i.

Given l in situations (a) and (b), we shall wish to consider only those values of m and n for which $l_{m,n} \leqslant l$. Any values of m and n such that this inequality is satisfied will then give an experiment which is *a priori* adequate. It will be a matter of practical convenience which among these are selected as the most economical. In the particular case where equal weight is to be given to m and n we shall wish to choose the smallest value of n consistent with $l_{n,n} \leqslant l$. Let m_c and n_c be the chosen values; then it may be asserted that each experiment-scheme of the set Σ_{m_c, n_c} is *a priori* adequate and we shall refer to each of them as *most economical* relative to those standards of economy which may have been defined. It will then be completely arbitrary which of the set Σ_{m_c, n_c} is used in the actual trial.

Our definition of the most economical rule is not completely satisfactory; consider the following example:

An urn contains two balls. We suppose that the form of an experiment which aims to ascertain the nature of these balls must be that of drawing n balls from the urn at random and with replacement. We suppose the necessary defined elements of the situation to be as follows:

I *uncertainty level:*	ε
	A both black
II *prior hypotheses:*	B both white
	C one black, one white
III *acceptable terminal statements:*	hypotheses above
IV *criterion of economy:*	smallness of n

We consider the following rule an n-fold sample. It may easily be represented in the observation-hypothesis diagram.

If all balls are black assert A.
If all balls are white assert B.
If some are black and some are white assert C.

The probability of incorrect assertion if A or B is true is 0 and if C is true is $\frac{1}{2^{n-1}}$. As long as we chose n so that $\frac{1}{2^{n-1}} \leqslant \varepsilon$, we shall have therefore prescribed an experiment which is *a priori* adequate. The most economical experiment according to our definition will then be that associated with what value of n is the smallest consistent with the relation $\frac{1}{2^{n-1}} \leqslant \varepsilon$.

Consider next the following rule which makes use of an accredited randomising device *after* the main experiment has been performed.

If some are black and some are white, assert C.
If all are white, assert C with probability γ and B with probability $(1-\gamma)$.
If all are black, assert C with probability γ and A with probability $(1-\gamma)$.

In each case, one hypothesis only is asserted to be true. The probability of incorrect assertion is then γ if A or B is true and $\frac{1}{2^{n-1}}(1-\gamma)$ if C is true. If we chose γ so that these risks are equal, that is to say $\gamma = \frac{1}{1+2^{n-1}}$, the overall chance of false assertion is $\frac{1}{1+2^{n-1}}$, whatever the true situation is. So if we chose n so that $\frac{1}{1+2^{n-1}} \leqslant \varepsilon$ we shall again have an experiment which is *a priori* adequate. But $\frac{1}{1+2^{n-1}} \leqslant \frac{1}{2^{n-1}}$ and so, if, for example, ε were given as $^1/_9$ the first rule would require five samples to be taken whereas the second would require only four. Our definition of *economy* does not therefore take account of every means by which sample size may be reduced. A similar improvement might be hoped for under some circumstances by the introduction of post-experimental randomisation into the therapeutic trial; but we may suppose that any resulting gain would be small. We accordingly adhere to our original definition.

8. The solubility of the problem.

It is necessary to justify our assertion that the problem we have stated in fact admits of a solution. We require the analogue of *Bernoulli's* theorem for this type of inverse situation.

Let us suppose that the number of cures in the treated group is a and in the untreated group is c in a particular trial. As before α, β, γ, δ, m, n, N will be used and the four types of subject in the trial-pool will be referred to as α, β, γ, δ.

The total number of ways in which the trial subjects may be allocated is $\binom{N}{n} = \frac{N!}{m!\,n!}$. We wish to evaluate the number of ways in which this allocation may be made consistent with the result (a, c).

We may classify this outcome as follows, the s in the last column being arbitrarily assignable:

Group	Outcome	Number	Type of Subject	Number of these Types
Treated	Cured	a	α or β	s, (a—s)
Treated	Not cured	m—a	γ or δ	(γ—n+c+β—a+s), (m+n—γ—c—β—s)
Untreated	Cured	c	α or δ	(α—s), (c—a+s)
Untreated	Not cured	n—c	γ or β	(n—c—β+a—s), (β—a+s)

Thus the treatment group comprises:

s of type α; (a—s) of type γ;
(γ—n+c+β—a+s) of type γ; (m+n—γ—c—β —s) of type δ.

The probability of this occurring is:

$$\frac{1}{\binom{m+n}{n}} \binom{\alpha}{s} \binom{\beta}{a-s} \binom{\gamma}{\gamma-n+c+\beta-a+s} \binom{\delta}{m+n-\gamma-c-\beta-s}$$

The probability of the result (a, c), given α, β, γ, δ is therefore:

$$\frac{1}{\binom{m+n}{n}} \sum_{s} \binom{\alpha}{s} \binom{\beta}{a-s} \binom{\gamma}{\gamma-n+c+\beta-a+s} \binom{\delta}{m+n-\gamma-c-\beta-s}^{1}$$

the summation being over all values of s which render meaningful the combinatorial expressions.

The joint probability generating function of a and c is therefore:

$$\frac{1}{\binom{m+n}{n}} \sum_{a,c,s} \binom{\alpha}{s} \binom{\beta}{a-s} \binom{\gamma}{\gamma-n+c+\beta-a+s} \binom{\delta}{m+n-\gamma-c-\beta-s} t_1^a t_2^c$$

the summation again being over all relevant values of the variables.

This expression may be verified as the coefficient of $y^m z^n$ in $\Psi(y, z, t_1, t_2)$ where

$$\Psi = \frac{m!\,n!}{N!} t_1^m t_2^n (1+y+z)^{\alpha} (1+y+t_2^{-1}z)^{\beta} (1+t_1^{-1}y+t_2^{-1}z)^{\gamma} (1+t_1^{-1}y+z)^{\delta} \ldots \text{(i)}$$

More generally, suppose we take without replacement r samples of $n_1, n_2 \ldots n_r$ members from a population of N individuals each of which has r potential responses, one corresponding to each sample into which an individual may fall. If the overall conceivable responses are λ in number we must suppose in the absence of other information that the population consists of individuals which may be distributed

[1] If we put $\beta = 0$, $\delta = 0$ this reduces to $\binom{\alpha}{c}\binom{\gamma}{m-a} / \binom{N}{m}$ the expression which underlies the *Irwin-Fisher* exact null hypothesis test, cf. *Irwin* [1935].

among λ^r classes. Let $A_{i_1}, \ldots, i_r$ be the number which would yield response i_1 if placed in the first sample, i_2 if placed in the second and so on.

$A_{i_1,\ldots,i_r}$ summed over all values of $i_1, \ldots, i_r$ equals N. The samples being drawn and the responses ascertained, let the results of the experiment be such that in the p^{th} sample a_{pk} subjects make the response numbered k, $\sum_{k=1}^{\lambda} a_{pk} = n_p$.

Then the joint p.g.f. of the a's, dummy variables t_{ij} being attached to the a_{ij}, is the coefficient of $\prod_{i=1}^{r} x_i^{n_i}$ in

$$\frac{1}{\binom{N}{n_1,\ldots,n_r}} \prod_{k_1 \ldots k_r = 1}^{\lambda} \left(1 + \sum_{i=1}^{r} t_{ik_i} x_i\right)^{A_{k_1} \ldots k_r}$$

(i) however is sufficient for our purposes. The means and variances of a and c are most conveniently obtained from the known values for moments of the hypergeometric distribution

$$E(a) = m\frac{\alpha+\beta}{N}; \; E(c) = n\frac{\alpha+\delta}{N};$$

$$\text{var}(a) = \frac{N-m}{N-1} m \frac{\alpha+\beta}{N} \cdot \frac{\gamma+\delta}{N};$$

$$\text{var}(c) = \frac{N-n}{N-1} n \frac{\alpha+\delta}{N} \cdot \frac{\beta+\gamma}{N}.$$

The covariance of a and c is the coefficient of $y^m z^n$ in

$$\left(\frac{\delta^2 \psi}{\delta t_1 \, \delta t_2}\right) t_1 = 1, t_2 = 1 \; ; \; \text{cov}(a, c) = \frac{mn}{N^2 (N-1)} \left\{(\gamma+\delta)(\beta+\gamma) - N\gamma\right\}$$

$$\text{Let } x = \frac{a}{m} - \frac{c}{n}; \text{ and } E(x) = \frac{\beta-\delta}{N} \text{ then}$$

$$\text{var}(x) = \frac{(\alpha+\beta)(\gamma+\delta)}{mN(N-1)} + \frac{(\alpha+\delta)(\beta+\gamma)}{nN(N-1)} - \frac{1}{N^2(N-1)} \left\{(\alpha+\gamma)(\beta+\delta) + 4\beta\delta\right\}$$

and we may note in passing that the same result is true when

$$\alpha + \beta + \gamma + \delta = m + n < N$$

Now $\frac{(\alpha+\beta)(\gamma+\delta)}{N^2}$ and $\frac{(\alpha+\delta)(\beta+\gamma)}{N^2}$ are maximal when $\alpha+\beta = \gamma+\delta = \frac{N}{2}$, $\alpha+\delta = \beta+\gamma = \frac{N}{2}$. $(\alpha+\gamma)(\beta+\delta)+4\beta\delta$ ises sentially non-negative and is therefore a minimum when $\beta = \delta = 0$ and therefore:

$$\text{max.}\left[\text{var}(x)\right] = \frac{N}{N-1}\left[\frac{1}{4m} + \frac{1}{4n}\right]$$

We may now prove the statement made in p. 328 as a formal theorem:

Given: (i) an uncertainty level ε.

(ii) a set of acceptable terminal statements comprising strictly overlapping interval estimates of $\frac{\beta-\delta}{m+n}$

(iii) a range of prior hypotheses for each value of $(m+n)$.

Then: It is possible to choose m and n and an appropriate exclusion rule, so that the experiment performed by taking from $(m+n)$ individuals m at random and placing these into the treated group, the others into the control group, is *a priori* adequate relative to the entities defined.

Proof: It will be sufficient to consider the case where the range of prior hypotheses covers all possible values of α, β, γ, δ and where the acceptable terminal statements comprise all interval estimates of the form

$$\vartheta \leqslant \frac{\beta-\delta}{m+n} \leqslant \vartheta + l;\ -1 \leqslant \vartheta \leqslant 1-l,\ \ l \text{ constant.}$$

The proof follows directly from our expression for the maximum variance of x and the application of *Chebychev's* theorem:

$$\Pr\left\{\left|x-\frac{\beta-\delta}{m+n}\right| \geqslant \sqrt{\frac{m+n}{m+n-1}\left(\frac{1}{4m}+\frac{1}{4n}\right)}\right\} \leqslant \frac{1}{k^2}$$

Putting $\frac{1}{k^2} = \varepsilon$ and choosing m and n so that

$$2k\sqrt{\frac{m+n}{m+n-1}\left(\frac{1}{4m}+\frac{1}{4n}\right)} \leqslant l$$

We have:

$$\Pr\left\{\left|x-\frac{\beta-\delta}{m+n}\right| \leqslant \frac{l}{2}\right\} \geqslant 1-\varepsilon \text{ and } \Pr\left\{x-\frac{l}{2} \leqslant \frac{\beta-\delta}{m+n} \leqslant x+\frac{l}{2}\right\} \geqslant 1-\varepsilon$$

This expression provides possible statements all of which are of the form required.

9. *Practical consequences in the single-stage therapeutic trial.*

We have put forward a logical foundation for the simplest form of therapeutic trial. There remains to explore the practical implications of the theory. It will be convenient to elaborate a rule based on the observed value of $x = \frac{a}{m} - \frac{c}{n}$. For illustrative purposes we shall suppose that such a rule belongs to the set $\Sigma_{m,n}$.

We require to define for each value of x a group of acceptable hypotheses defined in terms of $\beta-\delta/m+n$.

Given α, β, γ and δ, x may be taken to be distributed approximately normally with mean $\beta-\delta/m+n$ and variance

$$\frac{(\alpha+\beta)(\gamma+\delta)}{m(m+n)(m+n-1)}+\frac{(\alpha+\delta)(\beta+\gamma)}{n(m+n)(m+n-1)}-$$
$$-\frac{1}{(m+n)^2(m+n-1)}\left\{(\alpha+\gamma)(\beta+\delta)+4\beta\delta\right\}$$

Let $\varrho=\beta-\delta$, then $\gamma=m+n-\alpha-\varrho-2\delta$ and

$$\text{var }(x)\sim\frac{1}{N^3mn}\left\{\alpha(3mnN+m^3+n^3)+\delta(mnN+m^3+n^3)+n^2N\varrho-[n(\alpha+\varrho+\delta)+m(\alpha+\delta)^2]\right\}$$

Consider the case where there are equal numbers in the two groups, $m=n=N/2$.

$$\text{var }(x)\sim\frac{1}{n^2}\left\{\alpha+\frac{\delta}{2}+\frac{\varrho}{4}-\frac{(2\alpha+2\delta+\varrho)^2}{8n}\right\}$$

This is a maximum with respect to variations of α and δ

(i) $\varrho>0$; when $\delta=0$, $\alpha=n-\varrho/2$, when it equals $\frac{1}{n^2}\left(\frac{n}{2}-\frac{\varrho}{4}\right)$

(ii) $\varrho<0$; when $\delta=-\varrho$, $\alpha=n+\varrho/2$, when it equals $\frac{1}{n^2}\left(\frac{n}{2}+\frac{\varrho}{4}\right)$

$$\text{so }\left(x-\frac{\varrho}{2n}\right)^2\leqslant\frac{h^2}{n^2}\left[\frac{n}{2}-\left|\frac{\varrho}{4}\right|\right]$$

gives an approximate rejection region corresponding to each value of ϱ and an uncertainty level equal to the proportionate area outside (mean $\pm h\times$ s.d.) in the normal curve; and this will be so independently of the values of α and δ. An *a posteriori* interval estimate for ϱ is therefore given by solving the equation

$$\left(x-\frac{\varrho}{2n}\right)^2=\frac{h^2}{n^2}\left(\frac{n}{2}-\left|\frac{\varrho}{4}\right|\right)$$

It may be verified that this gives approximate interval estimates of $\varrho/2n$ as follows:

(i) $|x|>h\sqrt{\frac{n}{2}}$; $x\pm\frac{h}{\sqrt{2n}}\sqrt{1-\frac{|x|}{n}}$

(ii) $|x|<h\sqrt{\frac{n}{2}}$; $x+\frac{h}{\sqrt{2n}}\sqrt{1-\frac{x}{n}}$ and $x-\frac{h}{\sqrt{2n}}\sqrt{1+\frac{x}{n}}$

Consider first the set of prior hypotheses which embraces all possible values of α, β, γ, δ and the minimal set of terminal statements (a) (p. 328). (i) and (ii) may conveniently be used to accommodate this situation. The maximum length of interval occurs when $x=0$ and is equal to $h\sqrt{\frac{2}{n}}$. We must therefore take n as $2\frac{h^2}{l^2}$. The resulting rule may be summarised as:

Observation (x)	Assertion
$\frac{l}{2}-1 < x < 1-\frac{l}{2}$	$x-\frac{l}{2} > \frac{\beta-\delta}{2n} < x+\frac{l}{2}$
$x < \frac{l}{2}-1$	$-1 < \frac{\beta-\delta}{2n} < l-1$
$x > 1-\frac{l}{2}$	$1-l < \frac{\beta-\delta}{2n} < 1$

If, for example, we chose $l = {}^1/_{10}$ (10%) and an uncertainty level $\varepsilon = {}^1/_{20}$ (h approximately equal to two) n must be taken to be at least 800; that is to say, with equal numbers in control and treated groups the number of trial subjects must be at least 1,600.

Consider now prior hypotheses as above but with δ postulated as zero. The corresponding minimal set of terminal statements will be (b), p. 328. The maximal length of interval occurs approximately when $x = \frac{h}{\sqrt{2n}}$, when it is

$$h\sqrt{\frac{2}{n}}\sqrt{1-\frac{h}{n^{3/2}\sqrt{2}}}$$

In any practical situation n will be large, so that, as before, the second factor is negligible and we must take $l = h\sqrt{\frac{2}{n}}$. The resulting rule is:

Observation	Assertion
$\frac{l}{2} < x < 1-\frac{l}{2}$	$x-\frac{l}{2} < \frac{\beta}{2n} < x+\frac{l}{2}$
$x < \frac{l}{2}$	$0 < \frac{\beta}{2n} < l-1$
$x > \frac{l}{2}$	$1-l < \frac{\beta}{2n} < 1$

No material economy ensues on this limitation of the range of prior hypotheses.

The consequences of using a set of terminal statements which is not minimal may be briefly examined in the second of these two cases. Suppose that we use the set of *a posteriori* interval estimates, (i) and (ii), as our potential terminal statements. We know that $\beta \geqslant 0$; should the observed difference x be negative (as may well occur by chance) we shall be led to make a statement of apparent high

accuracy. Further, it is a possibility that we may be led to assert that $\beta = 0$, a statement we know could not reasonably be made as the result of any experiment carried out on a finite scale; and in the extreme case, whilst privately feeling confident that β is small, we shall be unable to make any statement whatsoever if we adhere to a strict interpretation of the prescribed rule.

10. Multiple-choice experiments.

The principles outlined above and in particular the importance of defining problems which admit of theoretical solution may be illustrated by an example widely discussed in statistical literature. For our purpose the discussion is suitably summarised in *Fisher* [1952, chapter 2] and *Neyman* [1950, p. 272 et seq.]. We give *Fisher*'s statement of the problem, and discuss his interpretation in the light of the concepts and by recourse to the terminology we have introduced in dealing with the therapeutic trial.

A lady declares that by tasting a cup of tea made with milk she can determine whether the milk or the tea was first added to the cup. *Fisher* proposes that the truth of her assertion should be tested by setting up an experiment in which she is allowed to taste eight cups, four mixed (she is told) in one way, four (she is told) in the other and presented to her in random order. The experiment as stated admits of the following interpretation.

Let us suppose it is legitimate to take only two prior hypotheses into account:

(a) The lady possesses no discriminating ability.
(b) In the circumstances defined she possesses perfect discriminating ability.

Each possible outcome of the experiment may conveniently be specified by the number of successes recorded in classifying those cups with milk added first. The schema of the experiment is then:

Number of successes .	0	1	2	3	4
No discrimination . .	$\frac{1}{70}$	$\frac{16}{70}$	$\frac{36}{70}$	$\frac{16}{70}$	$\frac{1}{70}$
Perfect discrimination	—	—	—	—	1

If ε is to be taken to be $\geqslant {}^{1}/_{70}$ the experimental rule marked by the heavy type satisfies all the requirements we have laid down.

The situation thus postulated however is trivial. *Fisher* remarks that the lady's claim will probably be

...not that she could draw the distinction with invariable certainty, but that, though sometimes mistaken, she would be right more often than not...

Once this possibility has been admitted we appear to have relinquished any hope of making the experiment as it stands *holonomic*. That is to say, there appears to be no way of adding to the list of prior hypotheses precisely-defined alternatives to the two extremes. Indeed one characteristic feature of the significance test as developed by *Karl Pearson* and *Fisher* is that it is intended primarily to deal with experiments whose designs are not holonomic. *Fisher* attempts to avoid the difficulty raised by the lady's objection to the experiment with 8 cups by suggesting experiments with 10, 12 or more cups so that (p. 21)

...using larger numbers . . . a significant result could be obtained with a still higher proportion of errors...

on the part of the lady. He is, however, unable in the last resort to give any criterion by which a mutually acceptable number of cups may be offered.

Neyman discusses this difficulty at length and draws the conclusion that the all-important question of how large the experiment should be in order to ensure a reasonable probability of success cannot be answered. He therefore introduces another form of experiment to test the lady's claim. This seems open to objections as strong or stronger than those which can be made to that of *Fisher*.

The latter introduces without definition (p. 22) the concept of a *degree of sensory perception*. This forms the basis of *Neyman*'s treatment. On successive days the lady is to be offered pairs of cups of the two types. The order on each day is to be random. It is then to be supposed that the lady possesses a degree of sensory perception p, being the probability of success in each trial. The trials are presumed to be independent. As a result of the experiment the lady is either to be rewarded for her ability or is to have her claim refuted.

Granted these assumptions the new experiment will be holonomic with an infinite number of prior hypotheses represented by $0 \leqslant p \leqslant 1$ or, if we are prepared to allow that the lady cannot be negatively discriminating, $.5 \leqslant p \leqslant 1$. The details of the design will now depend on what minimal list of acceptable terminal statements the experimenter will wish to assign. The set which in our approach appears to come nearest to *Neyman*'s treatment is that which comprises the two statements:

$$\begin{matrix} p \geqslant p_0 \\ p \leqslant p_1 \end{matrix}, \text{ where } \tfrac{1}{2} < p_0 < p_1$$

If in the event $p \geqslant p_0$ is asserted then the lady is rewarded; if $p \leqslant p_1$ is asserted her claim will be denied. The details of procedure, however, are not relevant in the present context.

The assumptions made are not consistent with the viewpoint we have adopted in the therapeutic trial. Difficulties at once spring from the concept of a *degree of sensory perception*. *Fisher* does not attempt to define it; and it would perhaps be fair to say that he regards the concept as essentially indefinable. For *Neyman* it is a postulate. When he introduces it he supposes (p. 272) that it results from

> ...identity of conditions and the complete independence of the *n* successive classification of pairs of cups of tea...

It is difficult to see that this can be meant seriously. If conditions were in fact identical what factor would intervene to make the lady's response vary? The only formal explanation is that in this respect the lady's sensory equipment operates in a fashion characteristic of the roulette board. Alternatively, following *Fisher's* treatment of other problems, we may identify *the lady as she is at the time of the experiment* with a conceptual lady out of whose infinitely protracted tea-tasting the experience of the experiment is regarded as a random sample. The idea may be attractive, but it carries with it an embarrassing consequence (if we pursue *Neyman*'s illustration). If the experiment demonstrates the phenomenon, it is the conceptual lady who must in fairness be rewarded and if not, it is the conceptual lady whose pretentions must be exposed.

For *Neyman*, however, the *p* appears to be real. The successive trials of the lady are specifically postulated as independent and she is axiomatically endowed with an unvarying *p*. It is difficult to imagine a justification which would obtain general acceptance among biologists and experimental psychologists for either of these postulates.

As an alternative we may examine *Neyman*'s own experiment from the viewpoint of the therapeutic trial model. We then offer pairs of cups on successive days for *n* days, the order on each day being decided by the tossing of a coin. There are 2^n ways in which the presentations may occur and 2^n potential responses to each of these sets. There are thus $(2^n)^{2^n}$ possibilities with respect to the relevant attribute of the lady in these circumstances. These are the hypotheses which the experiment will be regarded as designed to discriminate between. Two of these possibilities represent the lady as having (in these circumstances) perfect powers of discrimination, one of

these being that in which she inevitably allocates all cups wrongly. 2^n of them are *null hypotheses*, that is to say the lady will make the same judgement whatever orders the pairs are given. Terminal statements will be conveniently defined in terms of a system by which each hypothetical lady is given a score. We have to score such hypotheses as follows (n = 3), *M* indicating that the cup with milk added first is presented first and *T* the one with tea added first.

orders of presentation	M	M	M	M	T	T	T	T
	M	M	T	T	M	M	T	T
	M	T	M	T	M	T	M	T
potential assertion	M	T	M	M	T	T	T	M
	M	T	T	T	M	M	T	M
	M	M	T	M	M	T	T	T

A convenient scheme will be that in which the *total* potential correct assignments is regarded as the lady's score. The example above gives therefore a score of 3+0+2+2+3+3+2+1 = 16/24. The null hypothesis scores are 12/24. A little consideration will show that

(i) other systems of scoring will be as arbitrary or more so than this one.
(ii) no rule can be designed which, for large *n*, can make the length of the interval estimate of the lady's true score arbitrarily small.

We may here draw an analogy with experiments which have been performed in order to detect the presence of *weak telepathic power* using a randomising device such as the shuffling of cards in the hope that valid inferences could be secured. The above example of the lady tasting tea indicates that no such experiment can be satisfactorily designed if attention is confined to a single subject.

Accordingly we consider an alternative experiment. We suppose we have at our disposal *n* ladies. Each is presented with two cups in random order, conditions being precisely specified and as carefully controlled as possible. We may classify each lady according to her potential responses to the two ways of presentation. Let us suppose:

α are indiscriminating, i.e. would say the same in either case.
β are potentially right in both cases.
γ are potentially wrong in both cases.

$$\alpha+\beta+\gamma = n$$

Let the result of the experiment be that a correct decisions are made and c wrong decisions, $a+c = n$.

$$\text{Then } \Pr(a, c \mid \alpha, \beta, \gamma) = \Pr(a-\beta, c-\gamma \mid \alpha) = \binom{\alpha}{a-\beta}\frac{1}{2^\alpha}$$

So a is approximately normally distributed with mean $\frac{\alpha}{2}+\beta$ and variance $\frac{\alpha}{4}$.

$$\Pr\left[(a-\frac{\alpha}{2}-\beta)^2 > h^2\frac{\alpha}{4}\right] = \varepsilon$$

Where ε equals the proportionate area outside (mean $\pm h \times$ s.d.) in the normal curve.

Let $\varrho = \beta-\gamma$, which we shall wish to estimate.

Then:

$$\Pr\left[(a-\frac{\varrho+n}{2})^2 \geqslant \frac{nh^2}{4}\right] \leqslant \varepsilon$$

If acceptable terminal statements are interval estimates of fixed lengths the appropriate estimate is

$$\frac{\varrho}{n} = \left(\frac{2a}{n}-1\right) \pm \frac{h}{\sqrt{n}}$$

and n may be chosen to adjust the length suitably.

The situation may thus be resolved if our aim is analogous to those customarily entertained in scientific enquiry. The dilemma remains if we wish to invent a parlour game which is not arbitrarily favourable to one of the participants.

11. Conclusions—The hypothetical infinite population.

It is especially important to emphasise one particular conceptual difference between our approach and that which is widely accepted, because the numerical outcome of our investigation differs little from the result that we should have obtained, if prepared to invoke the customary *hypothtical infinite population*. It would be difficult to exaggerate the importance of this concept in modern statistical work especially that deriving from the schools of *von Mises* and *R. A. Fisher*. *Fisher*'s memoir on the *Mathematical Foundations of Theoretical Statistics* [1922], accepts the analytical function of theoretical statistics and defines the purpose of statistical methods as the reduction of quantities of data. He asserts that:

> This object is accomplished by constructing a hypothetical infinite population of which the actual data are regarded as constituting a random sample.

More forcefully in the Proceedings of the Cambridge Philosophical Society [1925] he writes:

> Any body of numerical observations or quantitative data thrown into a numerical form as frequencies may be interpreted as a random sample of some infinite hypothetical population of possible values.

Later in the same paper, and to dispel any remaining doubt about the essentially Platonic nature of the device, he states:

> . . . briefly the hypothetical population is the *conceptual resultant* of the conditions we study. (italics inserted)

The present approach to the therapeutic trial shows that in certain circumstances it is possible to devise techniques, which may be called statistical but do conform to the implicitly-accepted non-Platonic standards of the experimental method. An experimenter may observe the blood pressure of four dogs to rise following the injection of a solution of adrenaline. He may feel very strongly that such a result would occur with *all* dogs. He may even be allowed privately to suppose that the result would occur in the *conceptual resultant* of the dogs he has been concerned with. Nevertheless, he will present his observational records in such a way as to make clear that they are referable to the particular dogs on which he has experimented. The reader will have no excuse for concluding that they are referable to all dogs or to a singular hypothetical dog.

The *ultimate* aim of a therapeutic trial may well be to predict what would be the effect of substituting treatment B for treatment A; and the concept of the hypothetical infinite population may equally be regarded as sufficiently vague to accommodate an end which is as plausibly ill-defined as this. We have however maintained that the *primary* aim, and the only aim with which the investigator is concerned in the statistical treatment of his results, is to make an appraisal of the benefit which it might legitimately be supposed would have accrued to that group of subjects which has passed under his surveillance if all had been given treatment B rather than treatment A.

The statistician may reply that the entire intention of his techniques is to extend the range of induction beyond his material. If this is granted it follows that the proportions cured in each group of a therapeutic trial will be regarded as *observational records*[1] which have a value in their own right. They may be submitted accordingly to arbitrary analysis. An analytical approach to statistical methodology is granted and with it an inevitable susceptibility of any inferences to the jibe that statistics are able to prove anything. In the writer's view, the observational record in the controlled therapeutic trial is that particular one of the antecedently-specified acceptable terminal statements which the findings of the experiment allow us to make.

[1] Cf. *Woodger, J. H.*, Biology and Language, Cambridge 1952.

The ultimate responsibility for supposing that this result has relevance in a wider domain rests, as is true of more firmly established methods, with the consumer and not with the producer.

One conclusion we have drawn as the result of our investigation is that the number of test subjects required in the trial of a therapeutic method in a condition which fulfills our initial requirements is likely to be very large. It may be suggested that making the best of small numbers and the ensuing bad job forces us to fall back on the more customary methods of statistical procedure. Alternatively, we may conclude that the therapeutic trial carried out at a statistical level may have a more narrow range of applicability than its more enthusiastic exponents suggest.

The author is very deeply indebted to Professor *Lancelot Hogben*, F. R. S., for the discussions which have led to this approach.

Summary.

The type of situation is defined which may demand a statistical method in assessing the efficacy of a therapeutic measure.

To deal with such situations a theory is developed by analogy with ordinary experimental procedure.

This theory rejects the analytical approach to statistical inference and devises *a priori* rules of procedure in experimentation which is statistical in nature.

As a result it dispenses with the notion of a hypothetical infinite population and acknowledges that inferences must cover only those subjects coming within the surveillance of the investigator.

Résumé.

Définition du genre de situation dans laquelle il peut être nécessaire d'employer des méthodes statistiques pour pouvoir juger de l'efficacité d'une mesure thérapeutique.

Concernant ces situations on propose une théorie qui soit en analogie avec un procédé expérimental ordinaire.

Cette théorie rejette la voie analytique menant aux conclusions statistiques et fixe *a priori* des règles sur la manière de procéder dans des expériences de caractère statistique.

Il en résulte que l'auteur s'écarte de la notion d'une population infinie hypothétique et détermine que les inférences doivent se rapporter uniquement aux personnes examinées par la personne qui fait l'observation.

Zusammenfassung.

Es wurde der Typus eines Zustandes definiert, in dem die statistische Methode bei der Beurteilung der Wirkungsfähigkeit einer therapeutischen Maßnahme nötig werden kann. Mit Rücksicht auf solche Zustände wurde eine Theorie aufgestellt in Analogie zu einer gewöhnlichen experimentellen Durchführung.

Diese Theorie verwirft den analytischen Weg zu statistischen Schlußfolgerungen und weist a priori Regeln für die Durchführung bei der Auflegung eines Experimentes statistischer Natur an.

Als Folgerung hiervon nimmt der Verfasser Abstand von einer hypothetisch infiniten Bevölkerung, und stellt fest, daß die Schlußfolgerungen nur die Individuen decken müssen, die unter Beobachtung des Untersuchenden fallen.

REFERENCES

Carnap, R.: Logical Foundations of Probability, Routledge and Kegan Paul. 1950. – *Fisher, R. A.:* Phil. Trans. *222*, 309, 1921; Proc. Camb. phil. Soc. *22*, 700, 1925; The Design of Experiments, 6th ed. Oliver and Boyd, Edinburgh 1951. – *Greenwood, M.* and *G. U. Yule:* Proc. roy. Soc. Med. *8*, 113, 1915. – *Hogben, L.* and *R. Wrighton:* Brit. J. soc. Med. *6*, 89, 205, 1952. – *Irwin, J. O.:* Metron *12*, 83, 1935. – *Kendall, M. G.:* Biometrika *36*, 101, 1949. – *Keynes, J. M.:* A Treatise on Probability. Macmillan, London 1921. – *Neyman, J.:* Actualités sci. industr. *739*, 25, 1938; First Course in Probability and Statistics, Holt, New York 1950. – *Pearson, K.:* Phil. Mag. *50*, 157, 1900; Biometrika *8*, 250, 1911. – *Uspensky, J. V.:* Introduction to Mathematical Probability, McGraw-Hill, New York/London 1937. – *Wald, A.:* Statistical Decision Functions. John Wiley & Sons, New York 1950.